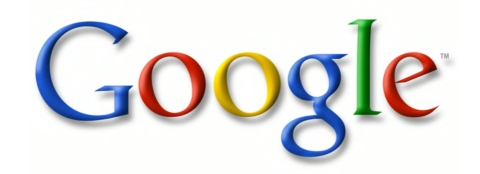
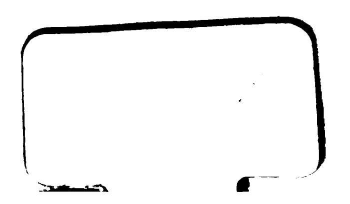




Die Kollektivmaßlehre

Von Hofrat E. Czuber

Professor der k. k. techn. Hochschule in Wien

Mit 5 Figuren im Text



Wien und Leipzig 1908 Kais. und kön. Hof-Buchdruckerei
und Hof-Verlags-Buchhandlung Carl Fromme

Sonderabdruck aus den „Mitteilungen über Gegenstände des Artillerie- und Geniewesens“,
Jahrgang 1907, zehntes Heft.

A.

Die Kollektivmaßlehre.

Nach einleitenden Vorlesungen von Hofrat Professor E. Czuber, gehalten an der k. k. Technischen Hochschule in Wien, zusammengestellt von Oberleutnant Alfred Strunz und Oberleutnant Karl Pietsch des k. u. k. Technischen Militärkomitees.

I. Einleitung.

Wenn wir die Naturobjekte, besonders jene organischer Natur, betrachten, so finden wir in ihren Merkmalen eine gewisse Stabilität. Diese bewirkt es, daß wir uns von jedem eine typische Vorstellung bilden. Diese Vorstellung wohnt uns inne, ohne daß wir sie genau präzisieren könnten.

Konkrete Beispiele mögen dies beleuchten. Wir besitzen eine typische Vorstellung von der Größe eines erwachsenen Menschen, ohne daß wir genau sagen könnten, worin diese Vorstellung besteht. Daß sie aber besteht, sehen wir daraus, daß wir nach ihr die Größe der Menschen beurteilen. Wir nennen einen Menschen klein oder groß, je nachdem er uns von der typischen Vorstellung, die wir uns gebildet haben, nach der einen oder anderen Seite abzuweichen scheint.

Starke Abweichungen von dem normalen Typus werden alle Menschen gleichartig konstatieren. Jeder wird einen Zwerg oder einen Riesen als solchen erkennen. Kleinere Abweichungen werden aber von verschiedenen Menschen verschieden beurteilt. Während der eine einen Menschen als groß bezeichnet, erkennt ihm der andere diese Eigenschaft noch nicht zu. Die typische Vorstellung ist daher bei verschiedenen Personen verschieden.

Ebenso haben wir uns eine typische Vorstellung von der menschlichen Hand gebildet, ohne daß wir die genauen Maße und Merkmale einer normalen Hand angeben könnten. Wir sprechen aber von einer schmalen oder einer breiten Hand, je nachdem das Verhältnis ihrer Länge zu ihrer Breite im Vergleich mit unserer Vorstellung von einer normalen Hand größer oder kleiner ist.

Weitere Beispiele werden die folgenden Betrachtungen bringen.

Einer der ersten, welcher Naturobjekte in Bezug auf ihre typische Erscheinung und die Abweichungen davon untersuchte, war der berühmte belgische Statistiker Quetelet. Er stellte diese Betrachtungen auch sofort auf den Boden, auf den sie gehören, auf den der Wahrscheinlichkeitsrechnung.

Wenn wir ein einzelnes Objekt, ein Exemplar, z. B. einen Schmetterling, einen Menschen u. s. w. auf seine Merkmale hin betrachten, seien diese nun qualitativer oder quantitativer Natur, so können wir dies einzelne Exemplar nur beschreiben, wie es z. B. in einem Steckbrief geschieht. Zu einer wissenschaftlichen Forschung, welche irgend welche Resultate allgemeiner Natur liefern könnte, bietet sich keine Grundlage. Diese ergibt sich erst bei Betrachtung einer Vielheit von Naturobjekten.

Zu dieser Betrachtung ist es notwendig, daß alle derselben unterzogenen Exemplare gewisse Bedingungen erfüllen.

1. Sie müssen eine gewisse Gleichartigkeit aufweisen.
2. Die Betrachtung darf sich nur auf ein Merkmal beziehen.
3. Das Merkmal muß sich durch Zahlen ausdrücken lassen.

Eine solche Vielheit von Naturobjekten, die sich nach einem Merkmal statistisch ordnen lassen, hat man unter dem Namen Kollektivgegenstand in die Wissenschaft eingeführt. Der Name Kollektivgegenstand bezeichnet daher immer eine größere Anzahl von Natur- oder auch Kunstgegenständen. Die Anzahl m der Exemplare, die zu einem Kollektivgegenstande vereinigt werden, nennt man seinen Umfang.

Einem solchen Kollektivgegenstande gegenüber stellt die Kollektivmaßlehre die Frage, ob und auf welche Weise er sich mathematisch beschreiben lasse.

Diese Namengebung und Fragestellung rühren von Theodor Fechner her, der mit vollem Recht als Begründer der Kollektivmaßlehre zu gelten hat, wenn auch Quetelet bereits ähnliche Fragen spezieller Natur gestellt hat.

Die Gleichartigkeit der Exemplare erfordert, daß sie sich hinsichtlich ihrer anderen Merkmale, auf welche sie nicht geprüft werden, entweder vollständig gleichen oder doch bis zu einem bestimmten Grade übereinstimmen. Der Grad dieser Übereinstimmung ist von großer Bedeutung; denn es ist klar, daß man schwer einen Kollektivgegenstand von größerem Umfange erhält, wenn man für die übrigen Eigenschaften große Übereinstimmung verlangt. Je tiefer man mit der Forderung der Gleichartigkeit geht, desto leichter wird man einen Kollektivgegenstand

von genügend großem Umfang finden. Fechner hat die Forderung aufgestellt, daß wenigstens die sogenannten Monstra, d. s. Objekte mit so abweichenden Merkmalen, daß sie aus dem typischen Rahmen offenbar herausfallen, ausgeschieden werden. Wenn man z. B. die Kopfumfänge von vielen Personen mißt, so vereinigt man diese Personen zu einem Kollektivgegenstand. Das Merkmal, nach welchem sie betrachtet werden, ist ihr Kopfumfang. Personen mit ganz ungewöhnlich großem oder kleinem Kopfumfang werden als Monstra ausgeschieden.

Die Ausscheidung der Monstra kann aber nicht schablonenmäßig erfolgen. Anders ist es, wenn man einen Kollektivgegenstand von nur wenig Exemplaren zur Verfügung hat, bei welchem die Einbeziehung der zufällig vorhandenen Monstra die Beobachtungsreihe über ihr Gewicht beeinflussen würde, anders, wenn ein Kollektivgegenstand von großem Umfange vorliegt. In diesem Falle ist es schwieriger zu entscheiden, ob man die Monstra, welche auf das Resultat der Beobachtung geringeren Einfluß üben, als zum Kollektivgegenstande gehörig betrachten soll oder nicht.

Ein weiteres Beispiel eines Kollektivgegenstandes, mit welchem sich Fechner viel beschäftigt hat, bildet die Körperlänge von Rekruten. Die Rekruten sind die Exemplare des Kollektivgegenstandes. Das Merkmal, auf welches hin sie betrachtet werden, ist ihre Körperlänge.

Die oben geforderte Gleichartigkeit in den übrigen Merkmalen zeigt sich hier in folgenden Umständen:

1. Ein bestimmtes Alter.
2. Gleiches Geschlecht.
3. Ein begrenzter Aushebungsbezirk, der eine Gleichartigkeit des Menschenschlages verbürgt.

Würde man Personen verschiedenen Alters, verschiedenen Geschlechtes, verschiedener Herkunft zu einem Kollektivgegenstande vereinigen, so erhielte man einen solchen von so ungleichartiger Beschaffenheit der Exemplare, daß die Untersuchung ihrer Körperlänge jede Bedeutung und jedes Interesse verlieren könnte.

Die Definition des Kollektivgegenstandes verlangt weiter, daß sich das Merkmal durch Zahlen ausdrücken lasse. Diese Bedingung ist immer erfüllt, wenn das Merkmal quantitativer Natur ist, daher durch Messung bestimmt werden kann. Bei manchen Kollektivgegenständen ist das Merkmal unmittelbar durch eine Zahl gegeben. Man zählt z. B. in einer Logarithmentafel auf 1 000 Seiten, wieviele Mantissen auf jeder

Seite mit 0 enden. Die Exemplare des Kollektivgegenstandes sind die einzelnen Seiten, das Merkmal, auf welches hin sie verglichen werden, bildet die Anzahl der Endnullen auf jeder Seite

II. Das Argument eines Kollektivgegenstandes. Stetige und unstetige Kollektivgegenstände.

Unter dem Argument eines Kollektivgegenstandes versteht man diejenige Größe, nach welcher die Exemplare statistisch geordnet werden. Es ist der zahlenmäßige, abgerundete Ausdruck für das betrachtete Merkmal, in der Folge konsequent mit x bezeichnet.

Das Argument ist daher für einen Kollektivgegenstand dasselbe, was die Variable für eine Funktion bedeutet.

Kann das Argument nur Werte annehmen, die nicht stetig aufeinanderfolgen, so nennt man den Kollektivgegenstand unstetig. In diesem Sinne ist der Kollektivgegenstand im letzten Beispiel unstetig, weil sein Argument, die Anzahl der Endnullen, eine unstetige Größe ist, die nur bestimmte Werte, ganze Zahlen, annehmen kann.

Wenn aber das Argument x wenigstens innerhalb bestimmter Grenzen, wenn auch nur theoretisch, aller Werte fähig ist, so bezeichnet man den Kollektivgegenstand als stetig.

Weil die mathematische Behandlung beider Fälle von einem gewissen Punkte an dieselbe ist, so werden sich die folgenden Betrachtungen nur auf stetige Kollektivgegenstände beschränken. Es ist übrigens nicht immer leicht zu entscheiden, ob ein vorliegender Kollektivgegenstand stetig oder unstetig ist.

Nennen wir z. B. die Anzahl der irgendwo vorkommenden Geburten g, die Zahl der männlichen m, die der weiblichen w. Es ist dann

$$g = m + w.$$

Das Verhältnis $\frac{m}{w} = x$ als Argument genommen scheint aller Werte fähig zu sein, da über die Anzahl g keine Annahme vorliegt.

Faßt man dagegen immer 100 Geburten zu einem Exemplare zusammen, so kann das Verhältnis $\frac{m}{g} = x'$ nur die Werte

$$0{,}00,\ 0{,}01,\ 0{,}02\ \ldots\ldots\ 1{,}00$$

annehmen, je nachdem unter den 100 Geburten keine, eine, zwei u. s. w. bis 100 männliche Geburten vorhanden sind. x' ist jetzt **unstetig**, es kann nur Werte zwischen 0 und 1, von Hundertel zu Hundertel steigend, annehmen.

Grenzen des Argumentes.

Ist das Argument des Kollektivgegenstandes unstetig, dann hat es mitunter von vornherein bestimmte Grenzen. Im obigen Beispiele ist die untere Grenze

$$u = 0,$$

die obere Grenze

$$o = 1;$$

im Beispiele der Abzählung der Endnullen auf den Seiten einer Logarithmentafel ist $u = 0$, wenn keine, $u = a$, wenn alle a Mantissen einer Seite mit 0 enden würden.

Bei stetigen Kollektivgegenständen, bis zu einem gewissen Grade auch bei unstetigen, lassen sich absolute Grenzen für das Argument nicht bestimmt angeben. Niemand kann z. B. scharfe Grenzen für die Körperlänge eines Menschen angeben. Betrachtet man einen Kollektivgegenstand von dem Umfange m, so wird man ein Exemplar mit der oberen Grenze o, eines mit der unteren Grenze u konstatieren, im obigen Beispiele den vorkommenden größten und kleinsten Menschen. o und u sind für diesen Fall die Grenzen des Argumentes.

Faßt man aus derselben Kategorie von Personen m' zu einem Kollektivgegenstande zusammen, so werden sich andere Grenzen des Argumentes o' und u' ergeben. Diese Grenzen werden wahrscheinlich weiter auseinander liegen, wenn $m' > m$ ist, enger sein, wenn $m' < m$ genommen wurde. Erst bei $m = \infty$ könnte man die wirklichen äußersten Grenzwerte feststellen. Dies ist aber praktisch unmöglich. Es unterliegt aber keinem Anstand, die Grenzen bis ∞ und $-\infty$ hinauszurücken, denn die sogenannten außermöglichen Fälle werden auch niemals konstatiert werden. Diese Annahme bringt daher für die analytische Behandlung der Kollektivgegenstände keinen Nachteil, ist aber für gewisse theoretische Folgerungen von Wichtigkeit.

Es gibt Kollektivgegenstände, die von mehreren Argumenten abhängig sind. Nehmen wir als Beispiel eine Anzahl von Eheschließungen. Die einzelne Ehe bildet das Exemplar. Die Ehen werden auf das Alter der Gatten bei der Eheschließung untersucht. Der Kollektivgegenstand besitzt also zwei Variable, das Alter des Mannes x und das Alter der Frau y.

Man kann auch hier den Kollektivgegenstand so behandeln, wie man sonst eine von zwei Variablen abhängige Funktion auf eine solche mit einer Variablen zurückführt. Man betrachtet zunächst alle Eheschließungen, bei welchen der männliche Teil ein bestimmtes Alter x_0 besitzt und forscht nach dem Alter der Frau. Alle diese Ehen bilden einen

Kollektivgegenstand für sich, abhängig von y. Durch Abstufung des x_0 löst man den Kollektivgegenstand mit zwei Variablen in eine Reihe von Kollektivgegenständen mit einer Variablen auf.

III. Erhebung eines Kollektivgegenstandes. Urliste.

Nach diesen Vorbegriffen wirft sich die Frage auf, was man erheben soll, um einen Kollektivgegenstand einer geregelten Forschung zu unterziehen.

Einen Kollektivgegenstand erheben heißt, für seine einzelnen Exemplare den Wert des ordnenden Argumentes feststellen. Das Erheben kann entweder aus dem bloßen Konstatieren des Wertes von x bestehen z. B. bei statistischen Erhebungen über das Alter von Personen auf Grund vorhandener Dokumente; oder in anderen Fällen in bloßem Abzählen, wie dies im Beispiel der Endnullen einer Logarithmentafel geschah. Ebenso hat man in neuerer Zeit vielfach als Kollektivgegenstand die Zahl der Blütenblätter zusammengesetzter Blüten benützt, da es sich zeigte, daß die Natur nur bei einfachen Blüten Regelmäßigkeit walten läßt. Auch hier bestand das Erheben des Kollektivgegenstandes in bloßem Abzählen.

Die wichtigste Art der Erhebung eines Kollektivgegenstandes bildet aber das Messen. Das Argument wird durch Messungen an den einzelnen Exemplaren erhalten: die Körperlänge, der Kopfumfang, das Gewicht des Gehirnes u. s. w.

Das Ergebnis der Erhebungen muß wie alle statistischen Erhebungen tabellarisiert werden. Diese einfache Tabelle nennt man die Urliste.

Schema einer Urliste.

Nummer des Exemplares	Am Exemplar konstatiertes x

Die Urliste läßt noch kein Gesetz erkennen, man wäre denn schon bei der Erhebung nach einer Regel vorgegangen. Dies wäre etwa der Fall, wenn man Messungen der Körperlänge von Personen nach deren annähernden Größe vorgenommen hätte.

IV. Die primäre Verteilungstafel.

Die Urliste muß erst weiter verarbeitet werden, ehe sie zu weiteren Untersuchungen benützt werden kann.

Es liege ein stetiger Kollektivgegenstand vor, dessen Erhebung durch Messen mit einem Maßstabe geschehen sei, z. B. Messen der Körperlänge. Der Maßstab hätte die Einheit $E = 1 cm$. Man konstatiert die Körperlänge immer in vollen Zentimetern. In diesem Falle werden die Exemplare, welche unter ein Argument x_0 zählen, d. h. eine bestimmte Körperlänge von x_0 Zentimeter aufweisen, nicht untereinander gleich sein. Ihre wirklichen Argumente bewegen sich zwischen den Grenzen $x_0 + \frac{1}{2} cm$ und $x_0 - \frac{1}{2} cm$, allgemein zwischen $x_0 + \frac{E}{2}$ und $x_0 - \frac{E}{2}$. Man faßt sie aber alle unter das Argument x_0 zusammen. Diesen Vorgang nennt man die Abrundung des Argumentes. Bei genauen Messungen wählt man E klein, bei ungenaueren groß.

Ordnet man die so zu Gruppen vereinigten Exemplare nach ihren Argumenten vom kleinsten beginnend in eine Tabelle, so erhält man die primäre Verteilungstafel, wie sie in ausführlicherer Form, als sie die Praxis verwendet, folgt:

Primäre Verteilungstafel.

Argumentwert x	Argumentintervall $x - \frac{E}{2}$ bis $x + \frac{E}{2}$	Anzahl der Exemplare z	Relative Häufigkeit der Exemplare y
x_0	$x_0 - \frac{E}{2}$ bis $x_0 + \frac{E}{2}$	z_0	y_0
.	.	.	.
.	.	.	.
.	.	.	.
x_i	$x_i - \frac{E}{2}$ bis $x_i + \frac{E}{2}$	z_i	y_i
.	.	.	.
.	.	.	.
.	.	.	.
x_g	$x_g - \frac{E}{2}$ bis $x_g + \frac{E}{2}$	z_g	y_g
	Summe . .	m	1

Nach dieser Tafel bewegt sich das Argument zwischen den Grenzwerten $x_0 - \frac{E}{2}$ und $x_g + \frac{E}{2}$. Aus praktischen Gründen geht man oft

über diese Grenzen, zwischen denen Exemplare beobachtet wurden, nach auf- und abwärts hinaus. Diese äußersten Intervalle, welche keine Exemplare enthalten, nennt man leere Intervalle.

Die dritte Kolonne der Tabelle sagt aus, wie viele Exemplare z_i erhoben worden sind, bei denen der Argumentwert x in das Intervall $x_i - \frac{E}{2}$ und $x_i + \frac{E}{2}$ fiel. Die Summe dieser Spalte ergibt m, die Anzahl der beobachteten Exemplare. Die Gesamtheit z_i der in ein Intervall fallenden Exemplare bezeichnet man als eine Variante des Kollektivgegenstandes. Jede Verteilungstafel enthält daher soviel Varianten, als Argumentwerte vorhanden sind.

Es empfiehlt sich immer, in die primäre Verteilungstafel noch die letzte Hilfsspalte y einzufügen, welche das Verhältnis $\frac{z_i}{m} = y$ enthält. Diese y_i sind relative Häufigkeiten im Sinne der Wahrscheinlichkeitsrechnung und haben daher die Bedeutung von empirischen Wahrscheinlichkeiten, d. h. sie geben die Wahrscheinlichkeit an, mit welcher ein zufällig herausgegriffenes Exemplar in die Variante z_i fällt, z. B. die Wahrscheinlichkeit, daß ein Rekrut 172 cm hoch ist.

x	z
x_1	z_1
.	.
.	.
.	.
.	.
x_i	z_1
.	.
.	.
.	.
.	.
.	.
.	.
x_n	z_n
	m

Nachdem die für die Verfassung der Verteilungstafel nötigen Daten und Begriffe festgelegt wurden, kann die praktisch übliche Zusammenstellung derselben erläutert werden.

Die primäre Verteilungstafel wird in der Regel zweispaltig angelegt, so zwar, daß man die Bildung der Intervalle dem Leser überläßt. Es sind zwei Formen der Tafel im Gebrauch:

1.) Die Verteilungstafel mit der absoluten Häufigkeit, welche die Argumente x_i und die Zahl der denselben entsprechenden Exemplare z_i, d. i. nichts anderes als die jeweiligen absoluten Häufigkeiten angibt. Da die Summe der letzteren m, d. i. dem Umfange des in Betracht gezogenen Kollektivgegenstandes gleich sein muß, läßt sich derart die Richtigkeit der Tafel überprüfen.

2.) Die Verteilungstafel mit der relativen Häufigkeit enthält die Aufeinanderfolge der veränderlichen Argumente und die zugehörigen relativen Häufigkeiten y_i, deren algebraische Summe stets 1 gibt.

x	y
x_1	y_1
.	.
.	.
.	.
.	.
x_i	y_i
.	.
.	.
.	.
.	.
.	.
.	.
x_n	y_n
	1

Es sei nun ein Beispiel einer solchen Verteilungstafel mit den absoluten Häufigkeiten über die Vertikalumfänge von Männerschädeln (gemessen von der Nasenwurzel über den Scheitel bis zum großen Hinterhauptsloch) angeführt. Als Maßeinheit wurde $E = 1mm$ gewählt, weshalb eine Variante stets alle Schädel umfassen wird, deren Umfänge bis $0{,}5mm$ größer oder kleiner sind als der Argumentwert.

x	z	x	z	x	z
368	1	397	7	419	9
371	2	398	14	420	15
376	1	399	3	421	8
378	1	400	13	422	7
379	1	401	12	423	5
380	2	402	13	424	12
381	1	403	6	425	8
382	2	404	10	426	7
383	3	405	18	427	3
384	3	406	8	428	4
385	8	407	8	430	3
386	2	408	16	431	3
387	6	409	13	432	2
388	4	410	20	433	5
389	5	411	9	434	5
390	7	412	15	435	4
391	7	413	8	438	1
392	7	414	12	440	3
393	2	415	21	442	1
394	8	416	6	443	1
395	12	417	5	447	1
396	4	418	16	448	1
					Summe: 450

Trotz der Unregelmäßigkeiten, welche die Tafel aufweist, kann man aus derselben entnehmen, daß die Werte z gegen die Mitte zunehmen, sich aber gegen die Enden hin verringern.

Wenn wir uns das Wesen der primären Verteilungstafel geometrisch versinnlichen, so finden wir das Gebiet der möglichen und eventuell auch jenes der außermöglichen Argumentwerte eingeteilt in eine Zahl Intervalle, die bei der vorhin angeführten Tafel gleich groß sind und deren Trennungspunkte man *Wechselpunkte* nennt. Jedem Intervalle kommt jener Argumentwert zu, welcher seiner Mitte entspricht, so daß, falls dieser mit x_i und die Intervallgröße mit E angenommen werden, der Argumentwert im linken Wechselpunkte $x_i - \frac{E}{2}$, im rechten aber $x_i + \frac{E}{2}$ beträgt.

$(x_i - 1)$ $\left(x_i - \frac{E}{2}\right)$ x_i $\left(x_i + \frac{E}{2}\right)$ $(x_i + 1)$

$\longleftarrow E \longrightarrow$

Die absolute Häufigkeit für dieses Intervall ist dann z_i, während die relative mit y_i bezeichnet werden möge.

Die Gleichheit der Intervalle wäre theoretisch nicht erforderlich, aber sie ist dennoch von der größten Wichtigkeit, da sie einerseits das Verständnis der ganzen Verteilungstafel wesentlich erleichtert, anderseits alle Rechnungsoperationen bei der Bearbeitung des Kollektivgegenstandes bedeutend vereinfacht, so daß ein weit umfangreicherer Apparat zur Bewältigung dieser Arbeit nötig würde, falls die Intervalle nicht gleich wären. Daher ist als oberster Grundsatz für die nähere Untersuchung eines Kollektivgegenstandes festzuhalten, daß man ihn vorerst auf eine Verteilungstafel von gleicher Intervallgröße bringe.

V. Reduzierte Verteilungstafel.

An der als Beispiel angeführten Verteilungstafel über die Vertikalumfänge von Männerschädeln ist wahrzunehmen, daß solche primäre Tafeln noch eine große Unregelmäßigkeit der absoluten bezw. relativen Häufigkeiten aufweisen.

Um das Verteilungsgesetz des Kollektivgegenstandes deutlicher hervortreten zu lassen, konstruiert man daher *reduzierte Verteilungstafeln*. Das Wesen einer solchen besteht darin, daß mehrere

Intervalle einer primären Verteilungstafel zu einem Intervalle der reduzierten zusammengezogen werden.

x_1 x_2 x_3 x_4 x_5 x_6 x_7 x_8

z_1 z_2 z_3 z_4 | z_5 z_6 z_7 z_8

$\frac{x_1 + x_4}{2}$ | $\frac{x_5 + x_8}{2}$

$z_1 + z_2 + z_3 + z_4$ | $z_5 + z_6 + z_7 + z_8$

Es sei in dem vorstehenden Schema einer primären Verteilungstafel x_1 das erste Intervall, in welchem sich z_1 Exemplare vorfinden und nun werde eine Verteilungstafel dadurch konstruiert, daß man immer vier Intervalle der primären zu einem Intervalle der reduzierten Tafel zusammenzieht. Da wirft sich sogleich die Frage auf, welche Beschreibung den neuen Intervallen zu geben ist.

Weil der erste in Betracht kommende Argumentwert $x_1 - \frac{E}{2}$ ist, die Größe des Argumentes am Schlusse des ersten Intervalls aber $x_4 + \frac{E}{2}$ beträgt,' so kommt der Intervallmitte der Wert des arithmetischen Mittels der Argumente der Wechselpunkte zu, d. i.

$$\frac{x_1 - \frac{E}{2} + x_4 + \frac{E}{2}}{2} = \frac{x_1 + x_4}{2}.$$

Dieser Betrag ist somit in die erste Spalte der neuangelegten Tafel einzusetzen, indes in die zweite die Summe der Exemplare, welche sich im zusammengezogenen Intervalle vorfinden, d. i. $z_1 + z_2 + z_3 + z_4$ einzuführen wäre, falls man mit absoluten Häufigkeiten zu arbeiten beabsichtigt oder $y_1 + y_2 + y_3 + y_4$, sobald die relativen benötigt werden.

Es ist ohneweiters klar, daß kein zwingender Grund besteht, die Reduktion unbedingt bei dem ersten Intervalle x_1 zu beginnen, sondern man kann den Anfangspunkt derselben in beliebiger Weise verlegen und mit der Reduktion bei x_2 oder x_3 oder x_4 einsetzen. Dadurch ergeben sich vier verschiedene reduzierte Tafeln und man spricht demgemäß von verschiedenen Reduktionslagen. Bei genauer Untersuchung zwingt sich die Erkenntnis auf, daß die Auswahl der Reduktionslage nicht gleichgültig ist; es zeigt häufig die eine dieser reduzierten Tafeln das Gesetz des Kollektivgegenstandes viel deutlicher als die andere.

Die Anlage einer solchen reduzierten Tafel unterscheidet sich von der primären nur durch die Größe der Intervalle. Zur Erläuterung sei eine aus der früher angeführten primären Tafel über die Schädelumfänge hervorgegangene reduzierte Verteilungstafel verzeichnet, bei welcher je fünf Intervalle zusammengefaßt wurden, so daß die neue Einheit $E = 5 \times 1mm = 5mm$ beträgt. Als unmittelbare Folge ergibt sich eine Verminderung der Zahl der Argumentwerte.

Argument x	Argumenteninterval $x - \frac{E}{2}$ bis $x + \frac{E}{2}$	Absolute Häufigkeit z
368	365,5 bis 370,5	1
373	370,5 „ 375,5	2
378	375,5 „ 380,5	5
383	380,5 „ 385,5	17
388	385,5 „ 390,5	24
393	390,5 „ 395,5	36
398	395,5 „ 400,5	41
403	400,5 „ 405,5	59
408	405,5 „ 410 5	65
413	410,5 „ 415,5	65
418	415,5 „ 420,5	51
423	420,5 „ 425,5	40
428	425,5 „ 430,5	17
433	430,5 „ 435,5	19
438	435,5 „ 440,5	4
443	440,5 „ 445,5	2
448	445,5 „ 450,5	2

Die Tafel ist in ausführlicherer Form, als dies gewöhnlich üblich, gegeben, indem die Intervallgrenzen angeführt werden.

Diese Tafel zeigt mit Ausnahme des Rückfalles von 17 auf 19 schon die charakteristischen Eigenschaften der reduzierten Tafeln, welche sich folgend aussprechen lassen:

Bei entsprechender Reduktion zeigen alle bisher untersuchten Kollektivgegenstände ein zunächst langsames, später rascher werdendes Ansteigen der absoluten Häufigkeiten, die Erreichung eines Maximums und von da ab ein langsames Abfallen gegen das Ende der Tafel; es tritt somit immer ein Berg ein.

Fechner war ursprünglich der Meinung, daß auch noch eine Symmetrie in der Verteilung stattfände, der Scheitel also in die Mitte der ganzen Tafel zu liegen käme. Die Abweichungen von dieser Annahme, die er konstatieren konnte, hielt er anfänglich nur für solche zufälliger Natur,

d. h. er hatte die Ansicht, es würden die Asymmetrien beim Vorhandensein hinlänglich vieler Exemplare verschwinden.

Weitere Untersuchungen aber haben gezeigt, daß die Symmetrie eigentlich eine Ausnahme ist und in der Regel Asymmetrien eintreten, welche bald größer, bald kleiner sind. In dem Aufsuchen des Unterschiedes in den Asymmetrien der verschiedenen Kollektivgegenstände liegt eben das Interessante der Studien über Kollektivmaßlehre.

Auch das eben angeführte Beispiel zeigt eine schwache Asymmetrie, indem der Berg mehr gegen das Ende der Tafel hin gerückt erscheint.

Interessante Asymmetrien ergeben sich bei Untersuchungen über die Häufigkeit des Auftretens gewisser Buchstaben in verschiedenen Sprachen. Wenn ein Druckwerk auf das Auftreten eines bestimmten Buchstaben in den Zeilen ausgezählt wird, so bekommt man einen Kollektivgegenstand. Die Zeilen sind dann die Exemplare, die Zahl der ausgezählten Zeilen ist der Umfang des Kollektivgegenstandes, der bisher immer mit m bezeichnet wurde, und die Wiederholungszahlen des betreffenden Buchstaben in den einzelnen Zeilen sind die Argumentwerte x.

So wurde ein klassisch geschriebenes Werk der deutschen Sprache, das uns also den höchsten Ausdruck der Sprache gibt, Kants „Naturgeschichte und Theorie des Himmels" in der jedermann leicht zugänglichen Ausgabe von Reclame in der angedeuteten Art untersucht; es erwies sich der Buchstabe „e" als der symmetrischeste Kollektivgegenstand, während das „t" nebenstehende unsymmetrische Verteilung zeigte.

x	z
0	56
1	208
2	294
3	244
4	114
5	55
6	23
7	3
8	3
	1 000

VI. Mathematische Charakterisierung eines Kollektivgegenstandes.

So wie man etwa Exemplare einer Spezies beschreiben kann, so ist es auch möglich daran zu denken, verschiedene Kollektivgegenstände zu beschreiben, um ihre Unterschiede und markanten Divergenzen klar zu charakterisieren. Es ist ein naheliegender Gedanke, der auch seit jeher praktiziert wurde, noch bevor man an eine Kollektivmaßlehre dachte, das arithmetische Mittel der Argumentwerte zur Charakterisierung des ganzen Gegenstandes zu verwenden, falls von einer Art eine große Zahl von Exemplaren vorliegt und diese auf ein quantitatives Merkmal geprüft werden.

Die Berechnung des arithmetischen Mittels, welches hier mit A bezeichnet sei, ergibt sich mit Hilfe der Verteilungstafel ohneweiters nach der Formel:

$$A = \frac{\Sigma\, x_i z_i}{\Sigma\, z_i}.$$

Da aber $\Sigma\, z_i$ dem Umfange der gesamten Tafel gleichkommt, so folgt die Vereinfachung:

$$A = \frac{\Sigma\, x_i z_i}{m} \quad \ldots\ldots\ldots\ldots \quad (1).$$

Man wird demnach von einem mittleren vertikalen Schädelumfang und von einer mittleren Häufigkeit der „t" sprechen.

Bedenkt man aber, daß fast alle Kollektivgegenstände eine Asymmetrie aufweisen, so ist sofort zu erkennen, daß das arithmetische Mittel A in der Kollektivmaßlehre nicht die erwünschte Charakteristik des Gegenstandes geben kann, ebenso wenig als etwa die Angabe eines „mittleren Gehaltes" einer Gruppe von Beamten Aufschluß über die tatsächlichen Bezüge eines einzelnen zu geben im stande ist. Daher empfiehlt es sich, zur Kennzeichnung des Kollektivgegenstandes noch andere Größen, sogenannte Hauptwerte, heranzuziehen.

Wir haben in jeder Verteilungstafel einen Berg konstatiert, dessen „Gipfel" in jenes Intervall fällt, welchem das größte z, d. i. die größte Anzahl von Exemplaren zukommt, welcher Umstand berechtigt, neben einem arithmetischen von einem dichtesten Mittel zu reden.

Das dichteste Mittel, welches mit D benannt werden soll, ist somit jener Argumentwert, welcher am häufigsten auftritt und gerade dieser Hauptwert scheint besonders charakteristisch für einen Kollektivgegenstand zu sein. Leider läßt sich für das dichteste Mittel keine einfache strikte Berechnungsregel aufstellen, seine Lage innerhalb des Intervalls mit dem größeren z nur durch einen Interpolationsprozeß mehr oder weniger verläßlich bestimmen.

Einen weiteren charakteristischen Hauptwert stellt der Zentralwert C des Kollektivgegenstandes dar, welcher derart zu liegen kommt, daß die Summen der absoluten Häufigkeiten oberhalb und unterhalb desselben einander gleich sind. Der Zentralwert hat auch eine leichtfaßliche wahrscheinlichkeitstheoretische Bedeutung, denn er bezeichnet jenen Argumentwert, der ebenso oft nicht erreicht als überschritten wird, welche Eigenschaft an den wahrscheinlichen Fehler erinnert.

Um den Zentralwert mathematisch zu charakterisieren, dient der Ansatz:

$$\sum_{x_1}^{C} z_i = \sum_{C}^{x_n} z_i \quad \ldots\ldots\ldots\ldots \quad (2).$$

Bei voller Symmetrie der Verteilung fallen alle drei bisher genannten Hauptwerte zusammen, je stärker aber die Asymmetrie wird, desto weiter rücken diese Werte auseinander, doch läßt sich auch theoretisch beweisen, daß der Zentralwert immer zwischen dem dichtesten Mittel und dem arithmetischen Mittel gelegen ist. Mit Rücksicht auf diese Tatsache erscheint auch die Bezeichnung Zentralwert sehr zutreffend gewählt.

Es empfiehlt sich immer bei einem Kollektivgegenstande die angeführten drei Werte zu bestimmen.

So liegt z. B. in der früher angegebenen Verteilungstafel der „t“ das dichteste Mittel im Intervalle 2, der Zentralwert dort, wo die Summe der absoluten Häufigkeiten $\frac{1000}{2} = 500$ erreicht wird, d. i. gegen das Ende des zu 2 gehörigen Intervalles, während das arithmetische Mittel, dem der Argumentwert 2,362 entspricht, in das zu 3 gehörige Intervall zu liegen kommt.

Auch die Angabe des Zentralwertes ist noch immer nicht ein genügend charakteristisches Merkmal für den Kollektivgegenstand. Einen viel genaueren Einblick gewährt die Konstruktion einer Größe, welche S t r e u u n g des Kollektivgegenstandes genannt wird. Diese Nomenklatur erscheint insoferne zweckmäßig, als sich die Exemplare über das Intervall zerstreuen und im gleichen Intervall bei dem einen Kollektivgegenstande sich mehr häufen als bei dem anderen.

Die Streuung stellt einen Begriff dar, welcher uns aus der Fehlertheorie rasch klar wird, da dort dieselbe Größe mit dem Ausdrucke mittlerer Fehler $\mu = \sqrt{\frac{[\lambda\lambda]}{n-1}}$ bezeichnet wird, worin λ die scheinbaren Fehler der einzelnen Beobachtungen, d. i. nichts anderes als die Abweichungen derselben vom arithmetischen Mittel und n die Zahl der Beobachtungen bedeuten. Nennt man die Streuung s, so kann sie analog wie der mittlere Fehler mit vollständig ausreichender Genauigkeit gerechnet werden nach der Formel:

$$s = \sqrt{\frac{\Sigma z_i (A - x_i)^2}{m}},$$

worin A das arithmetische Mittel, x_i den Argumentwert des betreffenden Intervalls und m den Umfang des Kollektivgegenstandes bezeichnen. Es hätte hier wenig Sinn, $m - 1$ in die Formel einzuführen, da ja stets eine sehr große Anzahl von Exemplaren betrachtet wird.

Wenn man nun zwei Kollektivgegenstände, welche sich selbstverständlich auf die gleiche Materie beziehen, vergleicht, also z. B. die Körper-

längen von Rekruten aus zwei verschiedenen Stellungsbezirken, wo für den einen dieser Kollektivgegenstände die Streuung s', für den anderen s'' gefunden wurde, so bedeutet dies, gleichen Umfang vorausgesetzt, wenn $s' > s''$, daß beim ersten Gegenstande die Exemplare auf eine längere Strecke verteilt sind als beim zweiten oder mit anderen Worten, daß der Berg im ersteren Falle niedriger und ausgedehnter als im zweiten ist. Es wäre hier wohl eine möglichst kleine Streuung erwünscht, weil dann die einzelnen Exemplare sich in ihrer Körperlänge weniger unterscheiden.

Hiemit sind alle Begriffe erläutert und jene Untersuchungen angedeutet, welche benötigt werden, um mit der eigentlichen, streng mathematischen Behandlung des Stoffes beginnen zu können; es erscheint daher das Ziel dieser Studie erreicht.

Zum Schlusse soll noch mit einigen Worten das weitere Streben der Kollektivmaßlehre angedeutet werden. Dieses geht natürlich über die Ermittlung des arithmetischen Mittels, des Zentralwertes etc. und der Streuung wesentlich hinaus und ist dahin gerichtet, den Kollektivgegenstand durch eine Funktion darzustellen und analytisch zu kennzeichnen, so daß es möglich wird, aus dieser analytischen Darstellung nachträglich theoretisch zu berechnen, wie viel Exemplare eines Kollektivgegenstandes von gegebenem Umfange auf ein bezeichnetes Intervall fallen. Man hat dann den gesamten Kollektivgegenstand in einer Formel dargestellt, ähnlich wie uns Kreis und Ellipse durch ihre Gleichungen gegeben sind; es ist tatsächlich bei verschiedenen Kollektivgegenständen schon in hohem Maße gelungen, dies zu erreichen.

Sonderabdruck aus den „Mitteilungen über Gegenstände des Artillerie- und Geniewesens“,
Jahrgang 1907, elftes Heft.

B.

Die Kollektivmaßlehre.

Von
Professor E. Czuber.

(Hiezu 5 Textfiguren.)

Die nachfolgenden Ausführungen schließen sich an die einleitenden Vorlesungen über die Kollektivmaßlehre an, welche in den Kapiteln I—VI niedergelegt sind, und setzen die daselbst entwickelten Grundbegriffe als bekannt voraus. Ihr Zweck ist, die mathematische Bearbeitung eines Kollektivgegenstandes bis zu einem Punkte zu führen, wo die praktische Anwendung in einem für viele Bedürfnisse, namentlich für vergleichende Studien ausreichenden Maße einsetzen kann.

I. Die Grundlage für das Studium und die mathematische Beschreibung eines Kollektivgegenstandes bildet eine aus dem Beobachtungsmateriale über denselben konstruierte Verteilungstafel. Auf die Struktur einer solchen Tafel muß nun näher eingegangen werden, wobei unterschieden werden soll, ob es sich um einen unstetigen oder einen stetigen Kollektivgegenstand handelt.

Bei einem unstetigen Gegenstande nimmt das Argument x nur eine diskrete Folge bestimmter Werte $x_1, x_2, \dots x_n$ an, von denen vorausgesetzt werden soll, daß sie nach einer arithmetischen Reihe fortschreiten, und noch spezieller, daß sie ganze um eine Einheit wachsende Zahlen bilden.

Zu jedem möglichen Argumentwert x_i hat nun die Beobachtung eine Zahl z_i ergeben, welche angibt, wie viele Exemplare dieses Argumentwertes gezählt worden sind. Zu den außermöglichen Argumentwerten gehört $z = 0$; indessen kann auch zu einem möglichen Argumentwert kein Exemplar vorgelegen haben, man spricht dann von leeren Argumentwerten.

Ist m der Umfang des Kollektivgegenstandes, so ist $\sum_1^n z_i = m$ und es bedeutet $\frac{z_i}{m}$ die relative Häufigkeit der Exemplare des Argumentwertes x_i oder, wie wir kürzer sagen wollen, des Argumentwertes x_i selbst. Setzt man $\frac{z_i}{m} = y_i$, so ist $\sum_1^n y_i = 1$.

Die Verteilungstafel enthält nun eine Kolonne der Werte x_i und eine zweite Kolonne mit den zugehörigen Zahlen z_i oder y_i. Ihr graphisches Bild, Fig. 1, besteht in einer Geraden XX' mit äquidistanten Punkten $x_1, x_2, \ldots x_n$, in welchen Ordinaten von den Längen $z_1, z_2, \ldots z_n$ oder den proportionierten $y_1, y_2, \ldots y_n$ (nach irgend einem ersichtlich zu machenden Maßstabe) aufgerichtet sind.

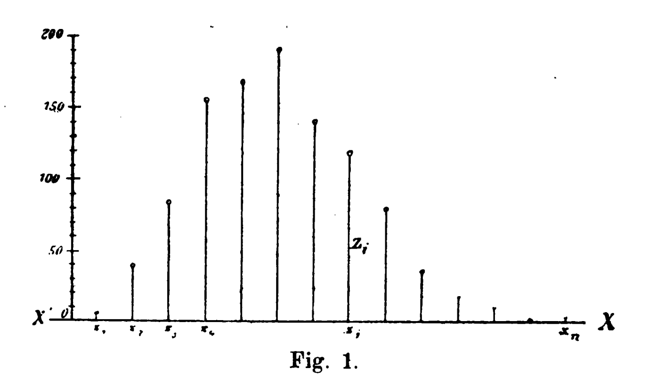


Fig. 1.

Wesentlich komplizierter liegen die Verhältnisse bei einem stetigen Kollektivgegenstand, bei dem das Argument x aller Werte eines gewissen zusammenhängenden Bereiches fähig ist. Aber schon die Beobachtung, die Messung, bestimmt den jeweiligen Argumentwert nicht genau, sondern mit einer Approximation oder Abrundung, was zur Folge hat, daß Exemplare, die in der Urliste mit demselben Argumentwert eingetragen sind, in Wirklichkeit innerhalb gewisser Grenzen verschiedene Argumentwerte besitzen. Häufig wird die kleinste Unterteilung E des benützten Maßstabes das Abrundungsintervall feststellen, indem Exemplare, deren Argument innerhalb der Grenzen $x - \frac{E}{2}$ und $x + \frac{E}{2}$ liegt, mit dem Argument x in die Liste eingesetzt werden.

Hat man auf Grund einer so gearteten Urliste eine primäre Verteilungstafel angefertigt, so schreiten sowohl die Argumentwerte $x_1, x_2, \ldots x_n$, als auch die Wechselpunkte $X_1, X_2, \ldots X_{n+1}$, die die Intervalle voneinander trennen, arithmetisch mit der Differenz E fort; unter dem Argument x_i sind alle Argumentwerte zusammengefaßt, die innerhalb der Grenzen X_i, X_{i+1} liegen.

Dementsprechend bedeuten die den Argumentwerten $x_1, x_2, \ldots x_n$ zugeordneten Zahlen $z_1, z_2, \ldots z_n$ oder die aus ihnen abgeleiteten $y_1, y_2, \ldots y_n$ nicht die absolute oder die relative Häufigkeit des Auftretens eines bestimmten Argumentwertes, sondern der Argumentwerte

eines Intervalls; insbesondere besagt das neben x_i stehende z_i bezw. y_i, daß in dem Intervall (X_i, X_{i+1}) z_i Exemplare beobachtet worden sind oder daß Exemplare dieses Ausmaßes mit der relativen Häufigkeit y_i vorkamen.

Die Verteilungstafel eines stetigen Kollektivgegenstandes und ihr geometrisches Bild, Fig. 2, unterscheiden sich von denen eines unstetigen dadurch, daß die Argumentwerte x_i Vertreter von Intervallen (X_i, X_{i+1}) sind und die z_i oder y_i sich nicht auf einzelne Argumentwerte, sondern auf eben diese Intervalle beziehen.

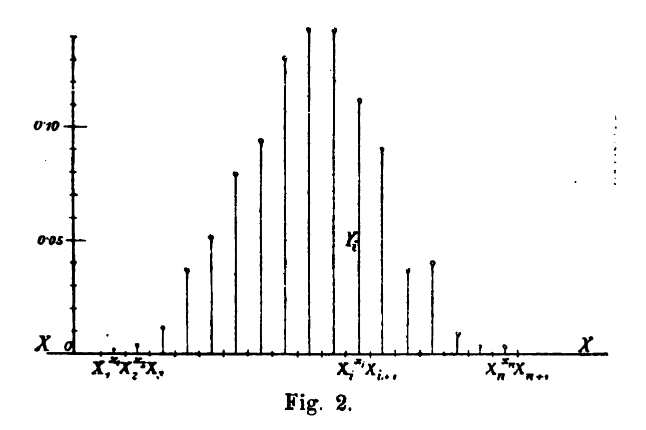


Fig. 2.

Man kann und wird übrigens bei der mathematischen Behandlung den Unterschied zwischen unstetigen und stetigen Kollektivgegenständen aufheben und die ersteren so behandeln wie die stetigen; es bedarf dann nur der Hinzufügung fingierter Wechselpunkte, welche in die Mitten zwischen den Argumentwerten verlegt werden.

Die Verteilungstafel hat, wie schon ihr Name andeutet, den Zweck, die Verteilung der Exemplare des Kollektivgegenstandes auf die Argumentwerte zur Anschauung zu bringen. Die primäre Tafel wird diesen Zweck vielfach nur in unvollkommener Weise erfüllen, umso unvollkommener, je ungünstiger das Verhältnis zwischen dem Umfang m und zwischen der Anzahl der ausgewiesenen Argumentwerte bezw. der Intervalle ist. Man wird dann zu einer reduzierten Tafel greifen, in der mehrere Intervalle der primären Tafel zu einem zusammengefaßt sind. Der Reduktion ist aber nach obenhin auch eine Grenze gesetzt. Leidet die primäre Tafel an dem Übelstand, daß zufällige Störungen den Gang der Verteilung nur undeutlich erkennen lassen, so kann wieder eine zu stark reduzierte Tafel charakteristische Einzelheiten der Verteilung verwischen. Es wird auf diesen Umstand nochmals hingewiesen werden.

II. Man stellt sich vor, daß einem wohldefinierten Kollektivgegenstand eine durch seine Natur bestimmte Verteilung zukommt, die dann zu genauem Ausdruck gelangen würde, wenn eine unbegrenzte (unendliche) Menge von Exemplaren zu Gebote stünde; ordnete man diese Exemplare nach ihren wirklichen Argumentwerten, so ergäbe sich die wahre Verteilung.

Die Aufgabe des mathematischen Teiles der Kollektivmaßlehre geht nun dahin, aus einer beschränkten Anzahl m von Exemplaren, richtiger, aus der über sie angefertigten Verteilungstafel, von der ideellen Verteilung so viel als möglich zu erfahren und festzustellen. Schon aus dieser Formulierung des Problems wird man erkennen, daß sich ein alleiniger wohl abgesteckter Weg zu seiner Lösung nicht wird angeben lassen.

Vor allem wird es notwendig sein, scharfe mathematische Begriffe einzuführen; die Grundlage soll der Begriff der Verteilungsfunktion bilden.

Die Verteilungsfunktion eines unstetigen Kollektivgegenstandes, $\mathfrak{U}(x)$, hat für die außermöglichen Argumentwerte den Wert 0 und drückt für die möglichen Argumentwerte $x_1, x_2, \ldots x_n$ die wahre relative Häufigkeit ihres Auftretens aus. Wird ein Kollektivgegenstand vom Umfang m beobachtet, so würden, wenn keine zufälligen Störungen stattfänden, $z_i = m\,\mathfrak{U}(x_i)$ Exemplare mit dem Argumentwert x_1 auftreten; zufällige Störungen bewirken, daß das beobachtete z_i von diesem ideellen abweicht; $\frac{z_i}{m} = y_i$ gibt den beobachteten Wert von $\mathfrak{U}(x_i)$. Im Sinne obiger Definition hat man sich $\mathfrak{U}(x)$ als Funktion einer unstetigen Variablen zu denken.

Bei einem stetigen Kollektivgegenstand ist die Verteilungsfunktion ein Grenzbegriff. Angenommen, bei m Exemplaren mögen sich Δz solche finden, deren Argument zwischen x und $x + \Delta x$, die Grenzen mit eingeschlossen, fällt; dann ist $\frac{\Delta z}{m}$ die relative Häufigkeit in dem Intervall $(x, x + \Delta x)$; sie selbst wird mit beständig wachsendem m der Grenze Null sich nähern, ihr Verhältnis zur Größe des Intervalls, d. i. $\frac{\frac{\Delta z}{m}}{\Delta x}$, wenn gleichzeitig Δx der Null sich nähert, wird aber einer bestimmten endlichen Grenze zustreben, die, weil von der Stelle x abhängig, eine Funktion $\mathfrak{V}(x)$ dieser Stelle ist und die relative Häufigkeit des Wertes x bedeutet. Dieser Definition zufolge ist $\mathfrak{V}(x)\,dx$ die relative Häufigkeit,

die dem Intervall $(x, x + dx)$ entspricht. Wird also ein Kollektivgegenstand vom Umfange m beobachtet, so sollten in das Intervall (X_i, X_{i+1}) zwischen zwei Wechselpunkten

$$z_i = m \int_{X_i}^{X_{i+1}} \mathfrak{B}(x)\, dx \quad \ldots \ldots \ldots \quad (1)$$

Exemplare fallen; zufällige Störungen aber bewirken, daß das beobachtete z_i von diesem ideellen abweicht.

Man nennt $\mathfrak{B}(x)$ die Verteilungsfunktion des stetigen Kollektivgegenstandes; ihre Bedeutung ist durch die Gleichung 1) und durch die aus ihr unmittelbar folgende

$$y_i = \int_{X_i}^{X_{i+1}} \mathfrak{B}(x)\, dx \quad \ldots \ldots \ldots \quad (2)$$

gekennzeichnet. Was ihren allgemeinen Verlauf betrifft, so wird ihr Wert im außermöglichen Gebiet von x durchwegs Null, im möglichen Gebiet durchwegs positiv sein; und bedeuten g, G die untere und obere Grenze dieses letzteren Gebietes, so ist

$$\int_g^G \mathfrak{B}(x)\, dx = 1;$$

statt dessen darf man, um alle Kollektivgegenstände zu umfassen, schreiben:

$$\int_{-\infty}^{\infty} \mathfrak{B}(x)\, dx = 1, \quad \ldots \ldots \ldots \quad (3)$$

da ja im außergewöhnlichen Gebiet $\mathfrak{B}(x) = 0$ ist.

III. Neben der Verteilungsfunktion spielt eine andere Funktion eine wichtige Rolle, deren Sinn und Bedeutung wieder zuerst für einen unstetigen, dann für einen stetigen Kollektivgegenstand erläutert werden soll.

Es seien

$$x_1, x_2, \ldots x_i, \ldots x_n \quad \ldots \ldots \ldots \quad (4)$$

die möglichen Argumentwerte bei einem unstetigen Kollektivgegenstande und

$$\mathfrak{U}(x_1), \mathfrak{U}(x_2), \ldots \mathfrak{U}(x_i), \ldots \mathfrak{U}(x_n)$$

die zugehörigen Werte seiner Verteilungsfunktion $\mathfrak{U}(x)$. Dann hat die Summe

$$\mathfrak{U}(x_1) + \mathfrak{U}(x_2) + \ldots + \mathfrak{U}(x_i) = \mathfrak{S}(x_i)$$

die Bedeutung der relativen Häufigkeit des Auftretens der Argumentwerte unter x_i mit Einschluß von x_i selbst. Mit Rücksicht darauf, daß $\mathfrak{U}(x)$ für außermögliche Werte von x Null ist, kann man sich die arithmetische Reihe (4) ins außermögliche Gebiet beiderseits ohne Ende,

bis $x_{-\infty}$ und x_{∞} fortgesetzt denken und dann $\mathfrak{S}(x_i)$ symbolisch in einer der folgenden Formen schreiben:

$$\mathfrak{S}(x_i) = \sum_1^i \mathfrak{U}(x_\nu) = \sum_{-\infty}^i \mathfrak{U}(x_\nu) \quad \ldots \ldots \ldots (5).$$

Streng genommen ist $\mathfrak{S}(x)$ wie $\mathfrak{U}(x)$ eine Funktion der unstetigen Variablen x_i ($i = 1, 2, \ldots n$). Man kann sie aber auch als unstetige Funktion der stetigen Variablen x auffassen; dann bleibt $\mathfrak{S}(x)$ zwischen zwei Werten x_i, x_{i+1} konstant $= \mathfrak{S}(x_i)$ und wächst beim Überschreiten von x_{i+1} unstetig um $\mathfrak{U}(x_{i+1})$. Bleibt man bei dieser Auffassung, so ist das geometrische Bild von $\mathfrak{S}(x)$, Fig. 3, im linken außermöglichen Gebiet die Achse $X'X$ selbst, im möglichen Gebiet eine stufenförmig ansteigende Folge von geraden Strecken parallel zu $X'X$, im rechten außermöglichen Gebiet eine zu $X'X$ parallele Gerade in der Entfernung 1.

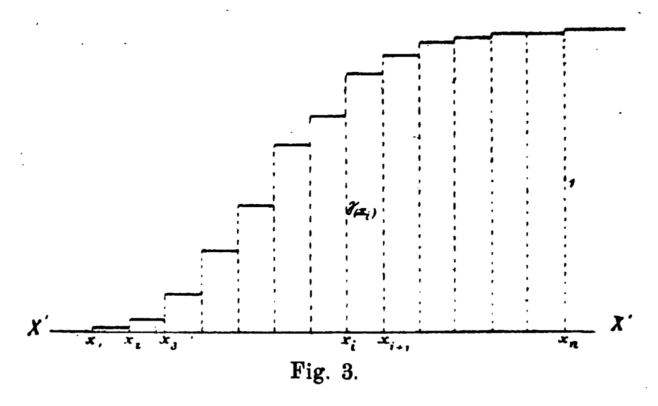


Fig. 3.

Ist $\mathfrak{V}(x)$ die Verteilungsfunktion eines stetigen Kollektivgegenstandes,

$$X_1,\ X_2,\ \ldots\ X_i,\ \ldots\ X_{n+1} \quad \ldots \ldots \ldots (6)$$

die Folge der äquidistanten Wechselpunkte im möglichen Gebiet von x, die unbeschadet auch bis $X_{-\infty}$ und X_{∞} fortgesetzt gedacht werden kann, so bedeutet wieder

$$\int_{X_1}^{X_2} \mathfrak{V}(x)\,dx + \int_{X_2}^{X_3} \mathfrak{V}(x)\,dx + \ldots + \int_{X_{i-1}}^{X_i} \mathfrak{V}(x)\,dx = \mathfrak{S}(X_i) \quad . \quad . (7)$$

die relative Häufigkeit der Argumentwerte unter X_i. Bezeichnet man allgemein X_i durch X, faßt die Integrale zu einem zusammen, und erstreckt mit Rücksicht darauf, daß $\mathfrak{V}(x)$ im außermöglichen Gebiet verschwindet, die untere Grenze bis $-\infty$, so hat man

$$\mathfrak{S}(X) = \int_{-\infty}^{X} \mathfrak{V}(x)\,dx \quad \ldots \ldots \ldots (8)$$

als Ausdruck der neuen Funktion, welche die relative Häufigkeit der Argumentwerte u n t e r X darstellt.

Das geometrische Bild dieser Funktion, die den Namen S u m m e n f u n k t i o n erhalten hat, wird im außermöglichen Gebiet links $X'X$ selbst sein; vom Eintritt in das mögliche Gebiet an steigt die Linie beständig und stetig und geht bei dem Übertritt ins außermögliche Gebiet rechts in eine zu $X'X$ im Abstande 1 parallele Gerade über (Fig. 4).

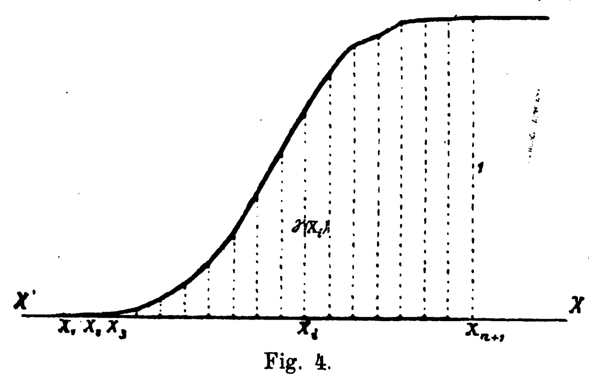


Fig. 4.

Der analystische Zusammenhang der Verteilungs- und der Summenfunktion ist nun folgender:

Kennt man von einem unstetigen Kollektivgegenstand die Verteilungsfunktion, also die Wertreihe $\mathfrak{U}(x_i)$ $(i = 1, 2, .. n)$, so bildet man die Summenfunktion nach der Vorschrift (5). Ist hingegen die Summenfunktion gegeben, d. h. die Wertreihe

$$\mathfrak{S}(x_1), \mathfrak{S}(x_2), \ldots \mathfrak{S}(x_i), \ldots \mathfrak{S}(x_n),$$

so braucht man nur die Differenzenreihe hiezu zu bilden, um die Verteilungsfunktion zu erhalten; denn

$$\mathfrak{U}(x_1) = \mathfrak{S}(x_1),\ \mathfrak{U}(x_2) = \mathfrak{S}(x_2) - \mathfrak{S}(x_1),\ \mathfrak{U}(x_3) = \mathfrak{S}(x_3) - \mathfrak{S}(x_2), \ldots$$

Weiß man die Verteilungsfunktion eines stetigen Kollektivgegenstandes, so findet man seine Summenfunktion nach der Vorschrift (8), also durch Integration; aus gegebener Summenfunktion hingegen ergibt sich die Verteilungsfunktion, indem man in (8) nach der oberen Grenze X differentiiert; es ergibt sich so

$$\mathfrak{V}(X) = \frac{d\,\mathfrak{S}(X)}{d\,X} \quad \ldots\ldots\ldots\ldots (9).$$

Um über die p r a k t i s c h e Bedeutung der beiden Funktionen das rechte Licht zu verbreiten, wird es notwendig sein, sie mit der Verteilungstafel in Beziehung zu setzen und zu fragen, welchen Aufschluß diese über die Verteilungs- und über die Summenfunktion bei einem unstetigen und einem stetigen Kollektivgegenstand zu geben im stande ist.

Diese Betrachtung wird am vorteilhaftesten an spezielle Fälle anzuknüpfen sein, die zugleich Anlaß geben zu mancherlei praktischen Bemerkungen.

Als Beispiel eines unstetigen Kollektivgegenstandes wähle ich die von H. Bruns[1]) ausgeführte Auszählung der ersten 1000 Spalten des Thesaurus logarithmorum Vegas auf die Endnullen; jede Spalte enthält 60 Mantissen. Exemplare des Kollektivgegenstandes sind die Spalten, sein Umfang daher $m = 1\,000$; Argument ist die Anzahl der in der Spalte angetroffenen Endnullen; das mögliche Gebiet des Arguments umfaßt die Zahlen von 0 bis 60, da selbst solche Spalten denkbar sind, die keine Endnull aufweisen und solche, die durchwegs mit einer Null endigen. Die Auszählung an 1 000 Spalten hat keine Spalte ohne Endnull und keine ergeben, in der mehr als 14 Endnullen vorkommen. Das Ergebnis ist in den ersten zwei Kolonnen von Tabelle I zusammengestellt. Die dritte Kolonne gibt in den y_i empirische Werte der Verteilungsfunktion $\mathfrak{U}(x)$ und in der vierten Kolonne empirische Werte der Summenfunktion $\mathfrak{S}(x)$. Man entnimmt der Tabelle, daß die relative Häufigkeit solcher Spalten, welche höchtens 7 Endnullen aufweisen, 0,747 beträgt, daß also nahe mit der Wahrscheinlichkeit $\frac{3}{4}$ zu erwarten ist, eine willkürlich herausgegriffene Spalte werde nicht mehr als 7 Endnullen zeigen.

Tabelle I.

x_i	z_i	$y_i = \frac{z_i}{m}$; $\mathfrak{U}(x_i)$	$\sum_1^i y_i$; $\mathfrak{S}(x_i)$
1	6	0,006	0,006
2	36	0,036	0,042
3	78	0,078	0,120
4	149	0,149	0,269
5	161	0,161	0,430
6	183	0,183	0,613
7	134	0,134	0,747
8	114	0,114	0,861
9	74	0,074	0,935
10	34	0,034	0,969
11	19	0,019	0,988
12	10	0,010	0,998
13	0	0,000	0,998
14	2	0,002	1,000
	1 000		

[1]) Wahrscheinlichkeitsrechnung und Kollektivmaßlehre (1906), pag. 256.

Der Fall eines stetigen Kollektivgegenstandes soll an der bereits in der Einleitung vorgeführten reduzierten Verteilungstafel der Vertikalumfänge von 450 Männerschädeln erklärt werden. Das Intervall, nach welchem die Argumentwerte und die Wechselpunkte fortschreiten, beträgt $\delta = 5 mm$. Die erste Spalte der Tabelle II führt das sogenannte Nummernargument, das von 1 zu 1 fortschreitet; das danebenstehende Argument x_i ergibt sich aus dem niedrigsten $x_1 = 368$ nach der Formel $x_i = x_1 + (i - 1)\delta$; die dritte Spalte enthält die Intervalle, die vierte die zugehörigen z_i. Die beiden nächsten Spalten geben die y_i und die Σy_i auf drei Dezimalen abgekürzt; infolge dieser Abkürzung schließt die Kolonne der Σy_i nicht genau mit 1, sondern mit 0,997. Da es jedoch für die fernere Bearbeitung von Wert ist, daß diese Kolonne mit 1 abschließe, so empfiehlt es sich, den umgekehrten Weg einzuschlagen und die Σy_i zuerst, die y_i aus ihnen durch Differenzbildung abzuleiten. Dieses Verfahren ist in den letzten drei Kolonnen zur Ausführung gebracht: erst sind die Summen $\sum_1^i z_i$, aus diesen dann durch Division mit $m = 450$ die Summen $\sum_1^i y_i$ und aus diesen durch Differenznehmen die y_i gebildet.

Es mag schon hier bemerkt werden, daß es sich empfiehlt, die Rechnungen soweit wie möglich mit den z zu führen; dadurch entgeht man den abgekürzten Divisionen.

Gegenüber dem früheren Falle besteht nun der wesentliche Unterschied darin, daß die y_i nicht empirische Werte der Verteilungsfunktion $\mathfrak{V}(x)$, sondern empirische Werte ihrer Integrale über die einzelnen Intervalle der dritten Kolonne bedeuten; hingegen liefern die Zahlen der Kolonne $\sum_1^i y_i$ empirische Werte der Summenfunktion $\mathfrak{S}(X)$.

Diese Betrachtungen lassen, soweit es sich um stetige Kollektivgegenstände handelt, eine Überlegenheit der Summenfunktion über die Verteilungsfunktion erkennen und legen den Gedanken nahe, die Aufsuchung der Summenfunktion als das primäre Problem zu betrachten; die Verteilungsfunktion ergibt sich dann als Resultat mathematischer Deduktion.

Tabelle II.

i	x_i (mm)	$X_i - X_{i+1}$	z_i	$y_i = \frac{z_i}{m}$	$\sum_1^i y_i = \mathfrak{S}(X_{i+1})$	$\sum_1^i z_i$	$\sum_1^i y_i = \frac{\sum_1^i z_i}{m} = \mathfrak{S}(X_{i+1})$	y_i
1	368	365.5 — 370,5	1	0,002	0,002	1	0,002	0,002
2	373	370,5 — 375,5	2	0,004	0,006	3	0,007	0,005
3	378	375,5 — 380,5	5	0,011	0,017	8	0,018	0,011
4	383	380,5 — 385,5	17	0,038	0,055	25	0,056	0,038
5	388	385,5 — 390,5	24	0,053	0,108	49	0,109	0,053
6	393	390,5 — 395,5	36	0,080	0,188	85	0,189	0,080
7	398	395,5 — 400,5	41	0,091	0,279	126	0,280	0,091
8	403	400,5 — 405,5	59	0,131	0,410	185	0,411	0 131
9	408	405,5 — 410,5	65	0,144	0,554	250	0,556	0,145
10	413	410,5 — 415,5	65	0,144	0,698	315	0,700	0,144
11	418	415,5 — 420,5	51	0,113	0,811	366	0,813	0,113
12	423	420,5 — 425,5	40	0,089	0,900	406	0,902	0,089
13	428	425,5 — 430,5	17	0,038	0,938	423	0,940	0,038
14	433	430.5 — 435,5	19	0,042	0,980	442	0,982	0.042
15	438	435,5 — 440,5	4	0,009	0,989	446	0,991	0,009
16	443	440,5 — 445,5	2	0,004	0,993	448	0,996	0,005
17	448	445,5 — 450,5	2	0,004	0,997	450	1,000	0,004
			450					

IV. Die ersten Ansätze zur Kollektivmaßlehre finden sich bei dem belgischen Statistiker A. Quetelet, in einer Schrift[1]), in der er sich die Aufgabe stellt, die Wahrscheinlichkeitstheorie auf die moralischen und politischen Wissenschaften zur Anwendung zu bringen. Methode und Material sind bemerkenswert genug, um ein kurzes Eingehen darauf zu rechtfertigen.

Für Quetelet stand es von vornherein fest, daß die Verteilung der Exemplare eines Kollektivgegenstandes dasselbe Gesetz befolge, wie die Verteilung der Abweichungen gleich genauer direkter Beobachtungen von ihrem arithmetischen Mittel, mit anderen Worten, die Exponentialfunktion des Gaußschen Fehlergesetzes galt ihm a priori als Verteilungsfunktion dermaßen, daß er eine Begründung dafür nicht für notwendig erachtete. Als Material standen ihm zu Gebote eine korrekt angelegte Verteilungstafel der Brustumfänge von 5 740 schottischen Soldaten und eine in ihren Enden stark reduzierte Verteilungstafel der Körperlängen von 100 000 französischen Rekruten.

[1]) *Lettres (à S. A. R. le Duc regnaut de Saxe-Cobourg et Gotha) sur la Théorie des Probabilités, appliquée aux Sciences morales et polytiques. Bruxelles, 1846.*

Bemerkenswert ist nun die Art, wie Quetelet den Anschluß der Verteilungstafel an die Funktion prüft. Er bedient sich dabei zweier Maßstäbe, die er als „Möglichkeitsmaßstab“ und als „Präzisionsmaßstab“ bezeichnet. Das Wesen dieser Maßstäbe ist mit der heute gebräuchlichen Terminologie so zu kennzeichnen: der Möglichkeitsmaßstab ist die Verteilungstafel, der Präzisionsmaßstab die Summentafel eines fiktiven unstetigen Kollektivgegenstandes, u. zw. besteht dieser aus den 1000 verschiedenen Kombinationen weißer und schwarzer Kugeln, die sich bei 999 Ziehungen ergeben können, vorausgesetzt, daß im einzelnen Zug der weißen und schwarzen Kugel je die Wahrscheinlichkeit $\frac{1}{2}$ zukommt. Die relativen Häufigkeiten oder Wahrscheinlichkeiten der einzelnen Kombinationen sind in aller Strenge durch die Glieder der Entwicklung $\left(\frac{1}{2}+\frac{1}{2}\right)^{999}$ gegeben. Bei dem großen Umfang ($m = 1000$) kann aber der Kollektivgegenstand angesehen werden wie ein stetiger mit der Verteilungsfunktion

$$\mathfrak{V}(x) = \frac{h}{\sqrt{\pi}} e^{-h^2 x^2} \quad \ldots\ldots\ldots\ldots \quad (10)$$

und der Summenfunktion

$$\mathfrak{S}(X) = \frac{1}{\sqrt{\pi}} \int_0^{hX} e^{-t^2} dt, \quad \ldots\ldots\ldots\ldots \quad (11)$$

beide Funktionen von jenem Wechselpunkte aus nach beiden Seiten gezählt, von dem aus die Verteilung symmetrisch angeordnet ist; dabei ist

$$h = \sqrt{\frac{2}{999}} = 0{,}0447437 \ldots\ldots\ldots\ldots (12).$$

Rechnet man nun $\mathfrak{S}(X)$ für

$$X = 1,\ 2,\ 3,\ 4, \ldots\ldots (500),$$

so bilden diese Werte in der angegebenen Reihenfolge Quetelets „Präzisionsmaßstab“ und die Differenzen

$$\mathfrak{S}(1) - 0,\ \mathfrak{S}(2) - \mathfrak{S}(1),\ \mathfrak{S}(3) - \mathfrak{S}(2), \ldots\ldots$$

seinen „Möglichkeitsmaßstab“; den aufeinanderfolgenden Zahlen beider Maßstäbe ordnet er die „Rangstufen“

$$1,\ 2,\ 3,\ 4, \ldots\ldots$$

zu. Anfang und Abschluß der Maßstäbe sind aus Tabelle III ersichtlich.

Tabelle III.

Kombination	Rangstufe	Möglichkeitsmaßstab	Präzisionsmaßstab
499 weiße u. 500 schwarze Kugeln oder umgekehrt	1	0,025225	0,025225
498 „ „ 501 „ „ „ „	2	0,025124	0,050349
497 „ „ 502 „ „ „ „	3	0,024924	0,075273
496 „ „ 503 „ „ „ „	4	0,024627	0,099900
495 „ „ 504 „ „ „ „	5	0,024236	0,124136
494 „ „ 505 „ „ „ „	6	0,023756	0,147892
493 „ „ 506 „ „ „ „	7	0,023193	0,171085
492 „ „ 507 „ „ „ „	8	0,022552	0,193637
491 „ „ 508 „ „ „ „	9	0,021842	0,215479
490 „ „ 509 „ „ „ „	10	0,021069	0,236548
. .	.	.	.
. .	.	.	.
. .	.	.	.
. .	.	.	.
420 weiße u. 579 schwarze Kugeln oder umgekehrt	80	0,00000004	0,4999992

Dieser Tabelle ist also beispielsweise zu entnehmen, daß bei einem stetigen Kollektivgegenstande, dem die Verteilungsfunktion (10) mit dem Parameterwert (12) zukommt, die Argumente der achten Stufe die relative Häufigkeit 0,022552, die Argumente der ersten acht Stufen die relative Häufigkeit 0,193637 besitzen; die Stufen sind dabei durch die Wechselpunkte 0, 1, 2, 3, 4,.... begrenzt, ausgehend vom Punkte der größten relativen Häufigkeit, in Bezug auf welchen Symmetrie in der Verteilung besteht.

Wie man sieht, führt die Tabelle von den 500 möglichen Stufen nur 80 an, weil die späteren wegen der außerordentlich geringen relativen Häufigkeit praktisch gar nicht in Betracht kommen.

Es bleibt noch zu erörtern, wie denn dieser eine Maßstab auf alle Kollektivgegenstände zur Anwendung gebracht werden kann, da doch der Parameter h, Gleichung (12), von dem Umfange abhängt. Diesen Punkt hat Quetelet[1]) mathematisch inkorrekt dargestellt, wenn auch die Schlußfolgerung, die er daraus für die Anwendung zieht, richtig ist. Der Grund und die Art der allgemeinen Anwendbarkeit beruht auf einer Transformation der Summenfunktion (11).

Soll der Maßstab, Tabelle III, auf einen Kollektivgegenstand von einem anderen Parameterwert h_1 angelegt werden, so wird die

[1]) l. c., pag. 105 bis 112.

Summenfunktion für die ersten X_1 Stufen dieses neuen Gegenstandes

$$\mathfrak{S}(X_1) = \frac{1}{\sqrt{\pi}} \int_0^{h_1 X_1} e^{-t^2} \, dt$$

lauten; soll nun $\mathfrak{S}(X_1) = \mathfrak{S}(X)$ sein, so muß notwendig

$$h_1 X_1 = h X$$

werden, woraus

$$X = \frac{h_1}{h} X_1; \quad \ldots\ldots\ldots\ldots \quad (13)$$

insbesondere entsprechen also einer Stufe des neuen Kollektivgegenstandes $\frac{h_1}{h}$ Stufen des Maßstabes. Da h, wie Formel (12) erkennen läßt, der Wurzel aus dem Umfange des Kollektivgegenstandes umgekehrt proportional ist, so verhalten sich äquivalente Stufenanzahlen wie die Quadratwurzeln aus den Umfängen der beiden verglichenen Gegenstände.

Nach dieser Vorbereitung wird es möglich sein, Quetelets Bearbeitung seines Beobachtungsmaterials verständlich zu machen.

Tabelle IV betrifft die Brustumfänge schottischer Soldaten. Die ersten drei Spalten bilden in aller Form eine nach gleichen Intervallen fortschreitende Verteilungstafel. In der vierten Spalte sind die z vom Umfang 5740 auf den Umfang 10000 durch proportionale Vergrößerung transformiert. Die beobachteten Wahrscheinlichkeiten der fünften Spalte sind in der Weise hergestellt, daß am oberen und unteren Ende von 0,5 ausgegangen und mit den Werten der vierten Spalte als Subtrahenden fortgeschritten ist (also $0{,}5 - 0{,}0005 = 0{,}4995$, $0{,}4995 - 0{,}0031 = 0{,}4964, \ldots$; ebenso von unten). Die zu diesen Wahrscheinlichkeiten gehörigen Rangstufen der nächsten Spalte, dem Präzisionsmaßstab entnommen, würden von der Mitte nach beiden Seiten in gleichen Intervallen fortschreiten, wenn die Verteilungstafel sich der supponierten Verteilungsfunktion genau anpaßte; dies ist jedoch nicht der Fall, und Quetelet nimmt auf Grund des Verlaufes dieser Zahlen an, daß einer Stufe der Verteilungstafel acht Stufen des Präzisionsmaßstabes entsprechen. Aus dieser Annahme ergaben sich die rechnungsmäßigen Rangstufen der nächsten Spalte, zu diesen die folgenden rechnungsmäßigen Wahrscheinlichkeiten nach dem Präzisionsmaßstab, schließlich durch Multiplikation mit 10000 und Differenzbildung von den beiden Enden gegen die Mitte die gerechnete Verteilung auf 10000 bezogen, die mit der beobachteten der vierten Spalte in Vergleich zu setzen ist.

Tabelle IV.

i	x_i engl. Zoll	z_i	z_i auf den Umfang 10000 bezogen	Beobachtete Wahrscheinlichkeit	Zugehörige Rangstufe	Rechnungsmäßige Rangstufe	Rechnungsmäßige Wahrscheinlichkeit	Rechnungsmäßige Verteilung
1	33	3	5	0,5000			0,5000	7
2	34	18	31	0,4995	52,0	50,0	0,4493	29
3	35	81	141	0,4964	42,5	42,5	0,4964	110
4	36	185	322	0,4823	33,5	34,5	0,4854	323
5	37	420	732	0,4501	26,0	26,5	0,4531	732
6	38	749	1305	0,3769	18,0	18,5	0,3799	1333
7	39	1075	1867	0,2464	10,5	10,5	0,2466	1838
				0,0597	2,5	2,5	0,0628	
8	40	1079	1882	0,1285	5,5	5,5	0,1359	1987
9	41	934	1628	0,2913	13,0	13,5	0,3034	1675
10	42	658	1148	0,4061	21,0	21,5	0,4130	1096
11	43	370	645	0,4706	30,0	29,5	0,4690	560
12	44	92	160	0,4866	35,0	37,5	0,4911	221
13	45	50	87	0,4953	41,0	45,5	0,4980	69
14	46	21	38	0,4991	49,5	53,5	0,4996	16
15	47	4	7	0,4998	56,0	61,5	0,4999	3
16	48	1	2	0,5000			0,5000	1
		5740	10 000					10 000

Das Beobachtungsmaterial über französische Rekrutenmaße, aus den ersten zwei Spalten von Tabelle V ersichtlich, ist nur in seinem zentralen Teile, von 1,570 bis 1,759*m*, nach Art einer Verteilungstafel, die in Intervallen von 27*mm* fortschreitet, gegliedert; am oberen Ende sind neun, am unteren sechs solche Intervalle zusammengefaßt. Quetelet hat nun aus dem gegliederten, über sieben Intervalle sich erstreckenden Mittelteil auf die ganze Verteilung geschlossen; die dabei befolgte Methode ist dieselbe, wie sie bei der vorigen Tabelle besprochen wurde, 5,8 Rangstufen sind als Äquivalent mit einem Intervall der Verteilungstafel angenommen worden.

So sehr ist Quetelet von der Geltung der Verteilungsfunktion (10) überzeugt, daß er die beträchtlichen Abweichungen zwischen rechnungsmäßiger und beobachteter Verteilung auf bestimmte Gründe zurückführen zu können glaubt. In den ersten neun Intervallen stehen 28 620 beobachteten Fällen 26 341 gerechnete gegenüber; das Plus der ersteren beträgt 2 279. Dieses Plus wird aber fast aufgewogen in den zwei nächsten Intervallen, wo 25 570 beobachteten Fällen 27 684 rechnungsmäßige gegenüberstehen. Quetelet zieht daraus den Schluß, daß aus

diesen Größenklassen viele zurückgestellt werden, die dann die Überfüllung der niedrigen Größenklassen verursachten. Die Assentkommission ziehe es eben vor, zu größeren Rekruten zu greifen, wenn sie ihr in genügender Zahl zur Verfügung stehen, statt solche zu behalten, die der unteren Grenze des Militärmaßes nahe sind.

Tabelle V.

i	X_i bis X_{i+1} (m)	z_i	Beobachtete Wahrscheinlichkeit	Zugehörige Rangstufe	Rechnungsmäßige Rangstufe	Rechnungsmäßige Wahrscheinlichkeit	Rechnungsmäßige Verteilung
1	1,327 — 1,354				62,2	0,49996	14
2	1,354 — 1,381				56,4	0,49982	47
3	1.381 — 1,408				50,6	0,49935	164
4	1,408 — 1,435				44,8	0,49771	449
5	1,435 — 1,462	28 620	0,50000		39,0	0,49322	1 105
6	1,462 — 1,489				33,2	0,48217	2 370
7	1,489 — 1,516				27,4	0,45847	4 440
8	1,516 — 1,543				21,6	0,41407	7 285
9	1,543 — 1,570				15,8	0,34122	10 467
10	1,570 — 1,597	11 580	5,21380	9,0	10,0	0,23655	13 182
11	1,597 — 1,624	13 990	0,09800	4,0	4,2	0,10473	14 502
			0,04190	1,6	1,6	0,04029	
12	1,624 — 1,651	14 410	0,18600	7,5	7,4	0,18011	13 982
13	1,651 — 1,678	11 410	0,30010	13,5	13,2	0,29814	11 803
14	1,678 — 1,705	8 780	0,38790	19,0	19,0	0,38539	8 725
15	1,705 — 1,732	5 530	0,44320	25,0	24,8	0,44166	5 627
16	1,732 — 1,759	3 190	0,47510	31,0	30,6	0,47355	3 189
17	1,759 — 1,786				36,4	0,48936	1 581
18	1,786 — 1,813				42,2	0,49621	685
19	1,813 — 1,840	2 490	0,50000		48,0	0,49881	260
20	1,840 — 1,867				53,8	0,49967	86
21	1,867 — 1,894				59,6	0,49993	26
22	1,894 — 1,921				65,4	0,49998	5
	über 1,921				71,2	0,50000	2
		100 000					99 996

V. Als G. Th. Fechner, der eigentliche Begründer der Kollektivmaßlehre, an den Aufbau derselben herantrat, scheint ihm zunächst das Gaußsche Exponentialgesetz auch als das prädestinierte Verteilungsgesetz von Kollektivgegenständen vorgeschwebt zu haben. Was über die Materie aus früherer Zeit vorlag, war dieser Annahme nicht ungünstig. Rekrutenmaße waren das erste Objekt, an dem das Gaußsche Gesetz erprobt wurde, und die in verschiedenen Ländern hierüber angestellten Unter-

suchungen ergaben durchwegs einen solchen Grad von Übereinstimmung, daß man glaubte in diesem Gesetze eine zureichende analytische Grundlage für die Darstellung der Verteilung gefunden zu haben.

Einen inneren Grund, der zur Anwendung des Gaußschen Gesetzes auf Kollektivgegenstände berechtigen würde, vermochte man nicht anzugeben. Fechner sprach es mit voller Bestimmtheit aus, daß die Abweichungen direkter Beobachtungen vom arithmetischen Mittel, für welche eben das Gaußsche Gesetz als geltend angesehen wird, aus ganz anderen Ursachen entspringen als die Abweichungen der Exemplare eines Kollektivgegenstandes von dem arithmetischen Mittel. Eine gewisse Analogie zwischen diesen beiden Dingen hat aber seinen Vorgängern und auch Quetelet vorgeschwebt: die Exemplare schienen durch den Zufall gestörte Verwirklichungen eines beabsichtigten, gewollten Modells oder Typus zu sein, ähnlich wie die einzelnen Messungen durch den Zufall gestörte Kopien des Originals darstellen.

In dem Maße nun, als Fechner den Kreis der untersuchten Kollektivgegenstände erweiterte, festigte sich ihm die Überzeugung, daß das Gaußsche Gesetz zur Darstellung aller nicht ausreiche. Es stellten sich nämlich bei manchen Gegenständen so beharrliche und starke Abweichungen von der Symmetrie, die eine so hervorstechende Eigenschaft jenes Gesetzes bildet, heraus, daß die Vorstellung, symmetrische Verteilung sei eine gesetzmäßige Eigenschaft aller Kollektivgegenstände, aufgegeben werden mußte. Vielmehr drang die Überzeugung durch, daß Asymmetrie verschiedenen Grades und verschiedener Richtung die Regel und volle Symmetrie nur eine seltene Ausnahme sei, wenn auch niedere Grade der Asymmetrie viel häufiger auftreten dürften als starke Abweichungen von der symmetrischen Anordnung.

Um diesen Tatsachen Rechnung zu tragen, ohne die rechnerischen Vorteile des Gaußschen Gesetzes aufgeben zu müssen, konstruierte Fechner das zweiseitige oder zweispaltige Gaußsche Gesetz, das, geometrisch gesprochen, aus zwei im Scheitel stetig ineinander übergehenden Kurvenästen besteht, deren jeder einer Gleichung von der Form

$$y = c\,e^{-h^2 x^2}$$

entspricht; doch sind die Parameter für die beiden Äste verschieden. (Fig. 5.)

Dieses Gesetz setzt die Existenz eines einzigen dichtesten Wertes, d. h. eines einzigen Argumentwertes D voraus, dessen relative Häufigkeit größer ist als die aller übrigen. Dieser dichteste Wert tritt insofern in die Rechte des arithmetischen Mittels ein, als er zum Ausgangswert für die Zählung der Abweichungen genommen werden muß; von ihm aus ist nach beiden Seiten getrennte Rechnung zu führen.

Angenommen, für einen (stetigen) Kollektivgegenstand, über den eine Verteilungstafel vorliegt, sei der dichteste Wert D bereits bestimmt; durch ihn zerfällt die Tafel in einen o b e r e n Teil (mit Argumentwerten $\overline{x}$ unter D) vom Teilumfang $\overline{m}$ und in einen u n t e r e n Teil (mit Argumentwerten $\underline{x}$ über D) vom Teilumfang $\underline{m}$, wobei $\overline{m} + \underline{m} = m$ wieder der ganze Umfang ist.

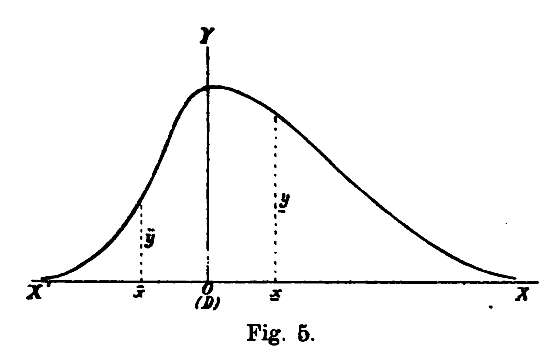


Fig. 5.

Sind $\overline{y}$, $\underline{y}$ die relativen Häufigkeiten der hervorgehobenen Argumentwerte, so sind

$$\overline{\Theta} = \Sigma \overline{y}\,(D - \overline{x}), \quad \underline{\Theta} = \Sigma \underline{y}\,(\underline{x} - D)$$

die d u r c h s c h n i t t l i c h e n Abweichungen im oberen bezw. unteren Teile; sie bilden neben $\overline{m}$, $\underline{m}$ diejenigen Größen, durch welche F e c h n e r die beiden Seiten charakterisiert. Die relative Häufigkeit der Argumentwerte zwischen zwei Grenzen $\overline{a}$, $\underline{a}$ ergibt sich dann durch den Ausdruck

$$\frac{1}{\sqrt{\pi}} \int_0^{\frac{\overline{a}}{\overline{\Theta}\sqrt{\pi}}} e^{-t^2}\, dt + \frac{1}{\sqrt{\pi}} \int_0^{\frac{\underline{a}}{\underline{\Theta}\sqrt{\pi}}} e^{-t^2}\, dt$$

und kann also mit Hilfe der vorhandenen Tafeln des Integralausdruckes $\frac{2}{\sqrt{\pi}} \int_0^t e^{-t^2}\, dt$ berechnet und sodann mit der beobachteten relativen Häufigkeit verglichen werden.

Die Lage des dichtesten Wertes ist unter der Geltung des zweispaltigen Gesetzes durch das von F e c h n e r so genannte P r o p o r t i o n a l i t ä t s g e s e t z gekennzeichnet, wonach er so liegt, daß

$$\overline{m} : \underline{m} = \overline{\Theta} : \underline{\Theta};$$

auf dieses Gesetz stützen sich die Methoden, die er ausgebildet hat, um die Lage von D in jenem Intervall der Verteilungstafel schärfer zu bestimmen, zu welchem das größte z gehört. Diese Bestimmung von D gehört unleugbar zu den Unbequemlichkeiten der ganzen Methode.

In dem Größenverhältnis von $\overline{m}$ und $\underline{m}$, von $\overline{\Theta}$ und $\underline{\Theta}$, dann in der Lage von D gegenüber dem arithmetischen Mittel A der Argumentwerte sprechen sich Grad und Richtung der Asymmetrie aus.

Ohne in das weitere Detail der von Fechner ausgebildeten Bearbeitungsmethode von Kollektivgegenständen einzugehen, soll zur Illustration des Vorgebrachten in Tabelle VI ein Beispiel vorgeführt werden. Es betrifft die Aufzeichnungen von 20 Jahrgängen Leipziger Studentenmaße. Die Bearbeitung ist einmal unter Zugrundelegung des einfachen Gaußschen Gesetzes, ein zweitesmal nach dem zweispaltigen Gesetz ausgeführt worden. Alle im Laufe der vorstehenden Betrachtungen genannten Elemente sind oberhalb der Tabelle mit ihren speziellen Werten angegeben.

Tabelle VI.

$A = 71{,}75;\quad D = 71{,}99;$

$m = 2047;\quad \overline{m} = 1083{,}5;\quad \underline{m} = 963{,}5;\quad \overline{\Theta} = 2{,}16;\quad \underline{\Theta} = 1{,}92.$

i	x_i (Zoll)	z_i	Rechnungsmäßige Verteilung (z_i)		Differenz zwischen der rechnungsmäßigen und der beobachteten Verteilung	
			nach dem einfachen Gaußschen Gesetze	nach dem zweispaltigen Gaußschen Gesetze	nach dem einfachen Gaußschen Gesetze	nach dem zweispaltigen Gaußschen Gesetze
1	60	1	—	—	— 1	— 1
2	61	0	—	—	0	0
3	62	0	—	0,5	0	+ 0,5
4	63	0	1	1,5	+ 1	+ 1,5
5	64	2	3,5	4	+ 1,5	+ 2
6	65	15,5	10	12	— 5,5	— 3,5
7	66	26	26	28	0	+ 2
8	67	54	58	59	+ 4	+ 5
9	68	108	110	108	+ 2	0
10	69	172	179	174	+ 7	+ 2
11	70	253	252	243	— 1	— 10
12	71	290	304	298	+ 14	+ 8
13	72	330,5	315	318	— 15,5	— 12,5
14	73	296	282	291	— 14	— 5
15	74	223,5	217	226	— 6,5	+ 2,5
16	75	142	143	145,5	+ 1	+ 3,5
17	76	75	81	80,5	+ 6	+ 5,5
18	77	38	40	37	+ 2	— 1
19	78	13	17	15	+ 4	+ 2
20	79	3,5	6	5	+ 2,5	+ 1,5
21	80	2	2	1	0	— 1
22	81	1	0,5	—	— 0,5	— 1
23	82	0,5	—	—	— 0.5	— 0.5
24	83	0,5	—	—	— 0,5	— 0,5
		2 047	2047	2047	abs. 90	abs. 72

Arithmetisches Mittel und dichtester Wert fallen beide in das Intervall 13 mit dem größten z, A ins dritte Viertel, D an das Ende; die Asymmetrie drückt sich in den Umfängen $\overline{m}$, $\underline{m}$ ziemlich deutlich aus; trotzdem ist der Anschluß an das einfache Gaußsche Gesetz nicht unbefriedigend, wenn auch der andere besser, wie dies insbesondere an den Summen der absoluten Werte der Differenzen zu erkennen ist.

Die Bruchzahlen in der Spalte z_i rühren daher, daß die Tafel keine primäre, sondern aus einer solchen, die von $^1/_4$ zu $^1/_4$ Zoll fortschritt, durch Reduktion entstanden ist.

VI. Es soll jetzt von der Möglichkeit einer analystischen Darstellung der Verteilungsfunktion oder der Summenfunktion abgesehen und die Frage so gestellt werden, welche Zahlwerte sich aus einer vorgelegten Verteilungstafel ableiten lassen, die geeignet wären, Schlüsse auf die Natur der Verteilung des betreffenden Kollektivgegenstandes zu gestatten, die insbesondere auch dazu dienen könnten, Kollektivreihen über ein und dieselbe Materie, die unter verschiedenen (örtlichen, zeitlichen) Umständen aufgenommen worden sind, auf objektiver Basis untereinander zu vergleichen. Solche charakteristische Größen sollen allgemein als Elemente des Kollektivgegenstandes bezeichnet werden.

Als geeignete Elemente sind gewisse auf das Argument gegründete Mittelwerte erkannt worden, da ja ihre Abhängigkeit von der Verteilung der Argumentwerte unmittelbar einleuchtet.

Angenommen, $f(x)$ sei eine endliche Funktion des Arguments, man bilde ihre Werte für alle Werte des Arguments nach Maßgabe ihrer Häufigkeit; dann ist der Durchschnitt dieser Werte der Mittelwert von $f(x)$ und soll mit $\mathfrak{D}[f(x)]$ bezeichnet werden. Nun ist $\mathfrak{V}(x)\,dx$ die relative Häufigkeit, $m\,\mathfrak{V}(x)\,dx$ die absolute Menge der Argumentwerte in dem Intervall $(x, x+dx)$, somit $m\,f(x)\,\mathfrak{V}(x)\,dx$ die auf dieses Intervall und $m\int_{-\infty}^{\infty} f(x)\,\mathfrak{V}(x)\,dx$ die auf das ganze Gebiet bezügliche Summe der Funktionswerte $f(x)$; ihre Division durch m gibt den Durchschnitt dieser Funktionswerte; also ist

$$\mathfrak{D}[f(x)] = \int_{-\infty}^{\infty} f(x)\,\mathfrak{V}(x)\,dx \quad \ldots \ldots \quad (14).$$

Diese Formel umfaßt alle erdenklichen Mittelwerte; einzelne davon werden durch Spezialisierung von $f(x)$ erhalten.

Vor allem werde die Wahl $f(x) = x$ getroffen; aus ihr geht das Mittel der Argumentwerte selbst, der **Argumentdurchschnitt** oder das arithmetische Mittel A hervor,

$$A = \mathfrak{D}(x) = \int_{-\infty}^{\infty} x\, \mathfrak{V}(x)\, dx \quad \ldots\ldots\ldots (15).$$

Die zu einer praktischen Ermittlung führende Formel ergibt sich wie folgt. Zerlegt man das Integrationsgebiet durch die Wechselpunkte der Verteilungstafel in Teilintervalle und ersetzt in einem solchen, wie (X_i, X_{i+1}), das variable x durch den mittleren Wert x_i, so tritt an die Stelle von $\int_{X_i}^{X_{i+1}} x\, \mathfrak{V}(x)\, dx$

$$x_i \int_{X_i}^{X_{i+1}} \mathfrak{V}(x)\, dx = x_i y_i = \frac{x_i z_i}{m},$$

an die Stelle von (15) also

$$A = \mathfrak{D}(x) = \frac{\sum_1^n x_i z_i}{m} \quad \ldots\ldots\ldots (16);$$

diese Formel gibt eine klare Vorschrift, wie aus der Verteilungstafel der empirische Wert von A abzuleiten ist.

Die Argumentwerte liegen nun algebraisch zu einem Teil unter, zum anderen Teil über dem arithmetischen Mittel A und zeigen von diesem teils negative, teils positive Abweichungen, die alle in dem allgemeinen Ausdruck $x - A$ enthalten sind; erhebt man diese Abweichungen zur Potenz p und wählt nun $f(x) = (x - A)^p$, so soll der Mittelwert hievon mit ε_p^p bezeichnet werden, so daß

$$\varepsilon_p^p = \mathfrak{D}[(x - A)^p] = \int_{-\infty}^{\infty} (x - A)^p\, \mathfrak{V}(x)\, dx \quad \ldots\ldots (17)$$

Die p^{te} Wurzel hieraus, also ε_p selbst, nennt man den aus den p^{ten} Potenzen abgeleiteten Mittelwert von $x - A$.

Es mag gleich bemerkt werden, daß der aus den ersten Potenzen gebildete Mittelwert theoretisch den Wert Null hat; denn es ist

$$\varepsilon_1 = \mathfrak{D}(x - A) = \int_{-\infty}^{\infty} (x - A)\, \mathfrak{V}(x)\, dx$$

$$= \int_{-\infty}^{\infty} x\, \mathfrak{V}(x)\, dx - A \int_{-\infty}^{\infty} \mathfrak{V}(x)\, dx = A - A = 0.$$

Für die praktische Durchführung der Formel (17) ergibt sich auf dem gleichen Wege, wie er im Anschlusse an die Formel (15) eingeschlagen worden ist, der Ansatz:

$$\varepsilon_p^p = \mathfrak{D}\left[(x-A)^p\right] = \frac{\sum_1^n (x_i - A)^p z_i}{m} \quad . \; . \; . \; . \; . \; (18).$$

Die unmittelbare Überlegung belehrt schon darüber, daß die Potenzmittel gewisse Rückschlüsse auf die Verteilung gestatten. Zunächst ist ohneweiters einzusehen, daß bei einer symmetrischen Verteilung alle Potenzmittel mit ungeradem Exponenten den theoretischen Wert Null haben, daß aber asymmetrische Verteilungen im allgemeinen durch von Null verschiedene Werte dieser Potenzmittel gekennzeichnet sein werden, wobei das Vorzeichen auf die Richtung der Asymmetrie hinweisen wird. Wenn also beispielsweise bei Kollektivreihen eines bestimmten Gegenstandes ε_3 stabil negativ ausfüllt, so deutet das auf eine asymmetrische Verteilung hin, welche Argumentwerte unter dem arithmetischen Mittel bevorzugt. Die Potenzmittel mit geraden Exponenten fallen naturgemäß durchwegs positiv aus. Wiewohl ihre Größe durch die ganze Verteilung beeinflußt wird, so gestattet doch die folgende Betrachtung einen allgemeinen Schluß. Wenn zwei Kollektivreihen sich derart voneinander unterscheiden, daß bei der einen die Glieder der Hauptmasse nach sich auf einem engen Argumentgebiet zusammendrängen, während sie bei der anderen über ein weiter ausgedehntes Gebiet verteilt sind, so werden bei der ersten die Potenzmittel mit geraden Exponenten im allgemeinen kleiner ausfallen als bei der zweiten. Es hängen also diese Potenzmittel mit der Ausbreitung des Kollektivgegenstandes in irgend einer Weise zusammen.

Hätte man auf Grund einer Verteilungstafel mit n Intervallen n Potenzmittel $\varepsilon_1, \varepsilon_2^2, \ldots \varepsilon_n^n$ berechnet, so wären diese im stande, die Verteilungstafel vollständig zu ersetzen, indem es dann möglich wäre, auf rein algebraischem Wege die relativen Häufigkeiten $y_1, y_2, \ldots y_n$ oder die absoluten $z_1, z_2, \ldots z_n$ u. zw. genau, wieder zu gewinnen. Eine kleinere Zahl von Potenzmitteln vermittelt eine bloß angenäherte Beschreibung, die aber für viele Fragen als ausreichend befunden werden kann. Jedenfalls gibt die Hinzufügung eines oder mehrerer Potenzmittel zum arithmetischen Mittel ein viel vollkommeneres Bild des Kollektivgegenstandes als die Angabe des arithmetischen Mittels allein, die bis in die jüngste Zeit häufig die einzige Charakterisierung abzugeben hatte.

VII. Um die technische Durchführung des eben auseinandergesetzten Verfahrens vorzubereiten, seien die maßgebenden Formeln (16) und (18) nochmals hiehergesetzt:

$$A = \mathfrak{D}(x) = \frac{\sum_1^n x_i z_i}{m}, \quad \varepsilon_p^p = \mathfrak{D}[(x_1 - A)^p] = \frac{\sum_i^n (x_i - A)^p z_i}{m}.$$

Die direkte Rechnung nach diesen Formeln würde schon bei einer Verteilungstafel von mäßiger Ausdehnung und selbst wenn man mit p nur etwa bis 2, 3, 4 gehen wollte, eine sehr beträchtliche Arbeit erfordern. Abgesehen von den zahlreichen Multiplikationen und Potenzierungen macht sich der Umstand als erschwerend geltend, daß A und hiemit auch die Differenzen $x_i - A$ vielstellige Zahlen sind.

Diesen Schwierigkeiten wird am wirksamsten dadurch begegnet, daß statt des arithmetischen Mittels ein anderer Ausgangswert u. zw. ein Argumentwert der Tafel selbst, gewählt wird. Dies hat zur Folge, daß dann die Abweichungen der Argumentwerte vom Ausgangswert Vielfache des Tafelintervalls sind, so daß der beschwerliche Teil der Rechnungen nicht mit den oft unbequemen Argumentwerten, sondern mit den ganzzahligen Nummernargumenten zu führen ist. Es handelt sich dann schließlich um den Übergang von dem (bis zu einem gewissen Grade willkürlich gewählten) Ausgangswert zu dem arithmetischen Mittel.

Der neue Ausgangswert heiße a und falle mit dem Tafelargument x_k zusammen, so daß

$$a = x_k \quad \ldots\ldots\ldots\ldots \quad (19).$$

Die auf diesen Ausgangswert bezogenen Potenzmittel der Abweichungen mögen mit η statt mit ε bezeichnet werden; es ist dann theoretisch

$$\eta_p^p = \mathfrak{D}[(x - a)^p] = \int_{-\infty}^{\infty} (x - a)^p \mathfrak{V}(x)\, dx \quad \ldots\ldots \quad (20),$$

während praktisch nach der Formel

$$\eta_p^p = \frac{\sum_1^n i\, (x_i - a)^p z_i}{m} \quad \ldots\ldots\ldots\ldots \quad (21)$$

zu rechnen sein wird.

Die folgende Tabelle VII zeigt die gegenwärtige Gestaltung der Abweichungen; an der Hand derselben wird sich die Struktur der Summen in (21) leicht überschauen lassen.

Tabelle VII.

Nummernargument i	Eigentliches Argument x_i	Abweichung vom Ausgangswert $x_i - a$	Absolute Häufigkeit von x_i z_i
1	x_1	$-(k-1)\delta$	z_1
2	x_2	$-(k-2)\delta$	z_2
.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.
$k-\nu$	$x_{k-\nu}$	$-\nu\delta$	$z_{k-\nu}$
.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.
$k-2$	x_{k-2}	-2δ	z_{k-2}
$k-1$	k_{k-1}	-1δ	z_{k-1}
k	$x_k = a$	0	z_k
$k+1$	x_{k+1}	1δ	z_{k+1}
$k+2$	x_{k+2}	2δ	z_{k+2}
.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.
$k+\mu$	$x_{k+\mu}$	$\mu\delta$	$z_{k+\mu}$
.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.
$n-2$	x_{n-2}	$(n-3)\delta$	z_{n-2}
$n-1$	x_{n-1}	$(n-2)\delta$	z_{n-1}
n	x_n	$(n-1)\delta$	z_n

Bei ungeradem p ist

$$\sum_1^n{}_i (x_i - a)^p z_i = -\delta^p \sum_1^{k-1}{}_\nu \nu^p z_{k-\nu} + \delta^p \sum_1^{n-k}{}_\mu \mu^p z_{k+\mu} \quad . \quad . \quad (22),$$

bei geradem p dagegen

$$\sum_1^n{}_i (x_i - a)^p z_i = \delta^p \sum_1^{k-1}{}_\nu \nu^p z_{k-\nu} + \delta^p \sum_1^{n-k}{}_\mu \mu^p z_{k+\mu} \quad . \quad . \quad . \quad (23);$$

es kommt also nur auf die Bildung der Summen

$$\sum_1^{k-1}{}_\nu \nu^p z_{k-\nu} = 1^p z_{k-1} + 2^p z_{k-2} + \ldots + (k-1)^p z_1 \quad . \quad . \quad (24),$$

$$\sum_1^{n-k}{}_\mu \mu^p z_{k+\mu} = 1^p z_{k+1} + 2^p z_{k+2} + \ldots + (n-k)^p z_n \quad . \quad . \quad (25)$$

an; diese aber lassen sich durch ein mechanisches Verfahren, das **nur** Additionen erfordert und daher als **Summenverfahren** bezeichnet wird, gewinnen.

Dieses Verfahren, zunächst auf den oberen Teil der Tabelle VII angewendet, besteht in folgendem. Man bildet nach und nach die **ersten Summen**:

$$\left.\begin{array}{llll} s_1^{(1)} & & & = z_1 \\ s_2^{(1)} & = s_1^{(1)} & + z_2 & = z_1 + z_2 \\ s_3^{(1)} & = s_2^{(1)} & + z_3 & = z_1 + z_2 + z_3 \\ \ldots & & & \\ s_{k-2}^{(1)} & = s_{k-3}^{(1)} & + z_{k-2} & = z_1 + z_2 + z_3 + \ldots + z_{k-2} \end{array}\right\} \quad . \; . \; (26),$$

die Summe dieser Summen heiße S_1^-, so ist

$$S_1^- = (k-2)\,z_1 + (k-3)\,z_2 + (k-4)\,z_3 + \ldots + z_{k-2};$$

fügt man die Summe

$$S_0^- = s_{k-2}^{(1)} + z_{n-1} = z_1 + z_2 + z_3 + \ldots + z_{k-2} + z_{k-1} \; . \; . \; (27)$$

hinzu, so ergibt sich

$$S_1^- + S_0^- = (k-1)\,z_1 + (k-2)\,z_2 + (k-3)\,z_3 + \ldots + 2\,z_{k-2} + z_{k-1},$$

d. h.

$$S_1^- + S_0^- = \sum_1^{k-1} \nu\, z_{k-\nu} \; . \; . \; . \; . \; . \; . \; . \; . \; . \; (28),$$

womit also die erste der Summen (24) gefunden ist.

Hierauf bildet man aus den ersten Summen (26) nach demselben Prinzip die **zweiten Summen**:

$$\left.\begin{array}{ll} s_1^{(2)} = s_1^{(1)} & = z_1 \\ s_2^{(2)} = s_1^{(2)} + s_2^{(1)} & = 2\,z_2 + z_2 \\ s_3^{(2)} = s_2^{(2)} + s_3^{(1)} & = 3\,z_1 + 2\,z_3 + z_3 \\ \ldots & \\ s_{k-3}^{(2)} = s_{k-4}^{(2)} + s_{k-3}^{(1)} & = (k-3)\,z_1 + (k-4)\,z_2 + (k-5)\,z_3 + \ldots + z_{k-3} \end{array}\right\} (29),$$

die Summe dieser Summen heiße S_2^-, also ist

$$S_2^- = \frac{(k-2)\,(k-3)}{2}\,z_1 + \frac{(k-3)\,(k-4)}{2}\,z_2 + \\ + \frac{(k-4)\,k-5)}{2}\,z_3 + \ldots + z_{k-3}$$

und

$$2\,S_2^- = (k-2)\,(k-3)\,z_1 + (k-3)\,(k-4)\,z_2 + \\ + (k-4)\,(k-5)\,z_3 + \ldots + 2\,z_{k-3};$$

das allgemeine Glied $(k-j-1)\,(k-j-2)\,z_j$ der rechten Seite kann aber entwickelt werden in

$$(k-j)^2\,z - 3\,(k-j)\,z_j + 2\,z\,,$$

infolgedessen wird

$$2 S_2^- = \sum_3^{k-1} \nu^2 z_{k-\nu} - 3 \sum_3^{k-1} \nu z_{k-\nu} + 2 \sum_3^{k-1} z_{k-\nu};$$

um nun die Summen rechts von 1 an laufen zu lassen, muß

$$2^2 z_{k-2} + 1^2 z_{k-1} - 3(2 z_{k-2} + 1 z_{k-1}) + 2(z_{k-2} + z_{k-1})$$

hinzugefügt werden; da sich dies aber auf Null reduziert, so ist auch

$$2 S_2^- = \sum_1^{k-1} \nu^2 z_{k-\nu} - 3 \sum_1^{k-1} \nu z_{k-\nu} + 2 \sum_1^{k-1} z_{k-\nu},$$

mit Rücksicht auf (28) und (27) also weiter

$$2 S_2^- = \sum_1^{k-1} \nu^2 z_{k-\nu} - 3(S_1^- + S_0^-) + 2 S_0^-,$$

woraus

$$2 S_2^- + 3 S_1^- + S_0^- = \sum_1^{k-1} \nu^2 z_{k-\nu} \quad \ldots\ldots\ldots \quad (30);$$

somit ist auch die zweite der Summen (24) gefunden.

In Fortsetzung des Verfahrens bilde man nun die dritten Summen:

$$\left.\begin{aligned} s_1^{(3)} &= s_1^{(2)} &&= z_1 \\ s_2^{(3)} &= s_1^{(3)} + s_2^{(2)} &&= 3 z_1 + z_2 \\ s_3^{(3)} &= s_2^{(3)} + s_3^{(2)} &&= 6 z_1 + 3 z_2 + z_3 \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ s_{k-4}^{(3)} &= s_{k-5}^{(3)} + s_{k-4}^{(2)} &&= \frac{(k-3)(k-4)}{2} z_1 + \\ &+ \frac{(k-4)(k-5)}{2} z_2 + \frac{(k-5)(k-6)}{2} z_3 + \ldots + z_{k-4} \end{aligned}\right\} (31),$$

die Summe dieser Summen heiße S_3^-, dann ist

$$S_3^- = \frac{(k-2)(k-3)(k-4)}{1.2.3} z_1 + \frac{(k-3)(k-4)(k-5)}{1.2.3} z_2 + \\ + \frac{(k-4)(k-5)(k-6)}{1.2.3} z_3 + \ldots + z_{k-4}$$

und

$$6 S_3^- = (k-2)(k-3)(k-4) z_1 + (k-3)(k-4)(k-5) z_2 + \\ + (k-4)(k-5)(k-6) z_3 + \ldots + 6 z_{k-4};$$

das allgemeine Glied der rechten Seite hat den Ausdruck $(k-j-1)$. $(k-j-2)(k-j-3) z_j$ und läßt sich entwickeln in

$$(k-j)^3 z_j - 6(k-j)^2 z_j + 11(k-j) z_j - 6 z;$$

dadurch wird

$$6 S_3^- = \sum_4^{k-1} \nu^3 z_{k-\nu} - 6 \sum_4^{k-1} \nu^2 z_{k-\nu} + 11 \sum_4^{k-1} \nu z_{k-\nu} - 6 \sum_4^{k-1} z_{k-\nu};$$

sollen aber die Summen rechts von 1 an laufen, so hat man hinzuzufügen

$$3^3 z_{k-3} + 2^3 z_{k-2} + 1^3 z_{k-1} - 6(3^2 z_{k-3} + 2^2 z_{k-2} + 1^2 z_{k-1}) + \\ + 11(3 z_{k-3} + 2 z_{k-2} + 1 z_{k-1}) - 6(z_{k-3} + z_{k-2} + z_{k-1});$$

da sich dies aber auf Null reduziert, so ist auch.

$$6 S_3^- = \sum_1^{k-1} \nu^3 z_{k-\nu} - 6 \sum_1^{k-1} \nu^2 z_{k-\nu} + 11 \sum_1^{k-1} \nu z_{k-\nu} - 6 \sum_1^{k-1} z_{k-\nu};$$

mit Benützung von (30), (28) und (27) gibt dies

$$6 S_3^- + 12 S_2^- + 7 S_1^- + S_0^- = \sum_1^{k-1} \nu^3 z_{k-\nu} \quad \ldots \ldots (32)$$

Der Weg der Ableitung ist jetzt deutlich vorgezeichnet, und es sei noch das Resultat des nächsten Ganges hier angegeben:

$$24 S_4^- + 60 S_3^- + 50 S_2^- + 15 S_1^- + S_0^- = \sum_1^{k-1} \nu^4 z_{k-\nu} \quad \ldots (33).$$

Die Summen (25) werden durch eine völlig analoge Rechnung im unteren Teile der Tabelle VII, welche von der untersten Zeile gegen den Ausgangswert hin zu führen ist, gefunden; die dabei auftretenden Hilfsgrößen sollen mit S_0^+, S_1^+, $S_2^+ \ldots$ bezeichnet werden.

Trägt man dann alles in die Formeln (22) und (23) ein, so ergeben sich wegen (21) die folgenden Endformeln für die Potenzmittel der Abweichungen vom Ausgangswert $a = x_k$, an deren Spitze die Formel für den Umfang der Kollektivreihe gestellt ist:

$$\left.\begin{aligned} m &= S_0^+ + z_k + S_0^- \\ m\eta_1 &= \delta\{(S_1^+ - S_1^-) + (S_0^+ - S_0^-)\} \\ m\eta_2^2 &= \delta^2\{2(S_2^+ + S_2^-) + 3(S_1^+ + S_1^-) + (S_0^+ + S_0^-)\} \\ m\eta_3^3 &= \delta^3\{6(S_3^+ - S_3^-) + 12(S_2^+ - S_2^-) + 7(S_1^+ - S_1^-) + (S_0^+ - S_0^-)\} \\ m\eta_4^4 &= \delta^4\{24(S_4^+ + S_4^-) + 60(S_3^+ + S_3^-) + 50(S_2^+ + S_2^-) + \\ &\quad + 15(S_1^+ + S_1^-) + (S_0^+ + S_0^-)\} \end{aligned}\right\} (34),$$

die sich noch etwas einfacher und übersichtlicher schreiben, wenn man sich der Abkürzungen

$$\left.\begin{aligned} S_i^+ + S_i^- &= \Sigma_i \\ S_i^+ - S_i^- &= \Delta_i \end{aligned}\right\} \ldots \ldots \ldots \ldots (35)$$

bedient, nämlich:

$$\left.\begin{aligned} m &= \Sigma_0 + z_k \\ m\eta_1 &= \delta(\Delta_1 + \Delta_0) \\ m\eta_2^2 &= \delta^2(2\Sigma_2 + 3\Sigma_1 + \Sigma_0) \\ m\eta_3^3 &= \delta^3(6\Delta_3 + 12\Delta_2 + 7\Delta_1 + \Delta_0) \\ m\eta_4^4 &= \delta^4(24\Sigma_4 + 60\Sigma_3 + 50\Sigma_2 + 15\Sigma_1 + \Sigma_0) \end{aligned}\right\} \ldots \ldots (36).$$

Die Bildung der Summen S^-, S^+, welche die Grundlage dieses Formelsystems bilden, geht nach dem aus der Tabelle VIII ersichtlichen Schema mechanisch vor sich.

Tabelle VIII.

i	x_i	z_i	$s_i^{(1)}$	$s_i^{(2)}$	$s_i^{(3)}$
1	x_1	z_1	$s_1^{(1)}$	$s_1^{(2)}$	$s_1^{(3)}$
2	x_2	z_2	$s_2^{(1)}$	$s_2^{(2)}$	$s_2^{(3)}$
3	x_3	z_3	$s_3^{(1)}$	$s_3^{(2)}$	. . .
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	$s_{k-4}^{(3)}$
$k-3$	x_{k-3}	z_{k-3}	$s_{k-3}^{(1)}$	$s_{k-3}^{(2)}$	S_3^-
$k-2$	x_{k-2}	z_{k-2}	$s_{k-2}^{(1)}$	S_2^-	
$k-1$	x_{k-1}	z_{k-1}	S_1^-		
k	$x_{k=a}$	z_k	.	.	.
$k+1$	x_{k+1}	z_{k+1}	S_1^+		
$k+2$	x_{k+2}	z_{k+2}	$s_{k+2}^{(1)}$	S_2^+	
$k+3$	x_{k+3}	z_{k+3}	$s_{k+3}^{(1)}$	$s_{k+3}^{(2)}$	S_3^+
.	.	.	.	.	$s_{k+4}^{(3)}$
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
$n-2$	x_{n-2}	z_{n-2}	$s_{n-2}^{(1)}$	$s_{n-2}^{(2)}$	.
$n-1$	x_{n-1}	z_{n-1}	$s_{n-1}^{(1)}$	$s_{n-1}^{(2)}$	$s_{n-1}^{(3)}$
n	x_n	z_n	$s_n^{(1)}$	$s_n^{(2)}$	$s_n^{(3)}$

Die Struktur dieses Schemas läßt sich kurz so beschreiben, daß von der vierten Kolonne an jede Zahl im oberen Teil gleich ist der über ihr stehenden vermehrt um die links neben ihr stehende, im unteren Teile gleich der unter ihr stehenden vermehrt um die links neben ihr stehende.

So umfassende Rechnungen bedürfen unerläßlich einer Kontrolle. Eine solche ergibt sich in sehr wirksamer Weise aus dem Aufbau der Systeme (26), (29), (31), . . . ; wenn man nämlich die letzten Gleichungen dieser Systeme paarweise addiert, so findet man:

$$s_{k-2}^{(1)} + s_{k-3}^{(2)} = S_1^-$$
$$s_{k-3}^{(2)} + s_{k-4}^{(3)} = S_2^- \quad . . . . \quad (37),$$
$$. . . . .$$

und ebenso ergibt sich im unteren Teile der Tafel:

$$s_{k+2}^{(1)} + s_{k+3}^{(2)} = S_1^+$$
$$s_{k+3}^{(2)} + s_{k+4}^{(3)} = S_2^+ \quad \ldots\ldots\ldots \quad (37^*).$$
. . . .

Man wird also vor dem Ziehen der Summe S_2^- (S_2^+) die Richtigkeit der Summe S_1^- (S_1^+), vor der Bildung von S_3^- (S_3^+) die Richtigkeit von S_2^- (S_2^+) u. s. w. prüfen.

Die Wahl des Ausgangswertes $x_k = a$ ist freigestellt; man kann ihn also auch in den letzten Argumentwert x_n verlegen oder in den ersten x_1. Tut man das erstere, so entfällt der untere Teil der Tabelle VIII und in den Formeln (34) verschwinden die S^+; entscheidet man sich für das zweite, so reduziert sich Tabelle VIII gewissermaßen auf ihren unteren Teil und in den Formeln (34) fallen die S^- aus. Bei Verteilungstafeln geringer Ausdehnung kann man tatsächlich einen dieser Wege einschlagen; bei ausgedehnteren Tafeln empfiehlt es sich aber, die Zweiteilung der Tabelle VIII eintreten zu lassen, indem man x_k etwa in der Mitte der Kolonne wählt; man erzielt damit den Vorteil, daß sich die ganze Rechnung in kleineren Zahlen bewegt.

Sind nun die Rechnungen bis zur Auswertung der Formeln (36), d. i. bis zur Bestimmung von

$$\eta_1, \eta_2^2, \eta_3^3, \eta_4^4, \ldots$$

gediehen, so handelt es sich um den Übergang zum arithmetischen Mittel, also zuerst um die Bestimmung dieses Mittels selbst und dann der Potenzmittel der Abweichungen von ihm, d. i. um

$$A; \varepsilon_1, \varepsilon_2^2, \varepsilon_3^3, \varepsilon_4^4, \ldots$$

Aus der allgemeinen Formel (20) ergibt sich für $p = 1$

$$\begin{aligned}\eta_1 = \mathfrak{D}[x-a] &= \int_{-\infty}^{\infty}(x-a)\,\mathfrak{W}(x)\,dx \\ &= \int_{-\infty}^{\infty} x\,\mathfrak{W}(x)\,dx - a\int_{-\infty}^{\infty}\mathfrak{W}(x)\,dx \\ &= A - a,\end{aligned}$$

folglich ist

$$A = \mathfrak{D}(x) = a + \eta_1 \quad \ldots\ldots\ldots \quad (38).$$

Die erste Frage ist also erledigt. Die zweite kann gleich für ein allgemeines p gelöst werden. Nach (18) ist

$$\varepsilon_p^p = \mathfrak{D}[(x-A)^p];$$

nun ist

$$x - A = \overline{x - a} - \eta_1,$$

daher

$$\begin{aligned}(x-A)^p = (x-a)^p - \tbinom{p}{1}(x-a)^{p-1}\eta_1 + \tbinom{p}{2}(x-a)^{p-2}\eta_1^2 - \ldots \\ \ldots + (-1)^{p-2}\tbinom{p}{2}(x-a)^2\eta_1^{p-2} + \\ + (-1)^{p-1}\tbinom{p}{1}(x-a)\eta_1^{p-1} + (-1)^p\eta_1^p;\end{aligned}$$

wendet man hierauf die $\mathfrak{D}$-Operation an, so entsteht zunächst

$$\varepsilon_p^p = \eta_p^p - \binom{p}{1} \eta_{p-1}^{p-1} \eta_1 + \binom{p}{2} \eta_{p-2}^{p-2} \eta_1^2 - \dots$$
$$\dots + (-1)^{p-2} \binom{p}{2} \eta_2^2 \eta_1{}^{p-2} + (-1)^{p-1} p\, \eta_1^p + (-1)^p \eta_1^p,$$

und schließlich, da sich die beiden letzten Glieder zusammenfassen lassen:

$$\varepsilon_p^p = \eta_p^p - \binom{p}{1} \eta_{p-1}^{p-1} \eta_1 + \binom{p}{2} \eta_{p-2}^{p-2} \eta_1^2 - \dots$$
$$+ (-1)^{p-2} \binom{p}{2} \eta_2^2 \eta_1^{p-2} + (-1)^{p-1} (p-1)\, \eta_1^p \quad . \; . \; . \; (39).$$

Gemäß dieser Formel hat man speziell:

$$\left.\begin{aligned} \varepsilon_1 &= 0 \\ \varepsilon_2^2 &= \eta_2^2 - \eta_1^2 \\ \varepsilon_3^3 &= \eta_3^3 - 3\,\eta_2^2\,\eta_1 + 2\,\eta_1^3 \\ \varepsilon_4^4 &= \eta_4^4 - 4\,\eta_3^3\,\eta_1 + 6\,\eta_2^2\,\eta_1^2 - 3\,\eta_1^4 \end{aligned}\right\} \; . \; . \; . \; . \; . \; . \; . \; (40).$$

VIII. Das Minimum an Elementen, durch welche man einen Kollektivgegenstand mathematisch beschreiben kann, bilden der Argumentdurchschnitt oder das arithmetische Mittel A der Argumentwerte und die Streuung, worunter die Quadratwurzel aus dem Durchschnitt der Abweichungsquadrate, also jene Größe zu verstehen ist, welche bisher mit ε_2 bezeichnet worden ist.

Der Argumentdurchschnitt ist nicht so sehr an sich, als vielmehr durch den Umstand von Bedeutung, daß er bei allen bisher untersuchten Kollektivreihen in die Nähe des dichtesten Wertes, also in jenen Teil des Argumentgebietes fiel, wo sich die Exemplare am dichtesten zusammendrängen. Wenn man also in dem arithmetischen Mittel auch nicht, wie dies früher vielfach geschah, den Ausdruck eines Typus, ein von der Natur gewolltes Modell zu erblicken hat, so bezeichnet es doch die Stelle, in deren Umgebung die Häufung der Exemplare größer ist als im ganzen übrigen Verlaufe, ist also eine eminent charakteristische Größe.

Die Streuung ε_2 ist, wie schon früher erklärt wurde, ein Ausbreitungsmaß. Nach ihrer Definition entspricht sie dem mittleren Fehler einer Beobachtungsreihe und in demselben Sinne, wie der dreifache mittlere Fehler als eine begründete Annahme für den größten absoluten Fehlerbetrag gelten kann, darf man auch die sechsfache Streuung als rundes Maß der Ausbreitung des Kollektivgegenstandes ansehen, also erwarten, daß die Hauptmasse der Exemplare in ein Argumentgebiet von der Länge der sechsfachen Streuung fallen werde.

Es geben also in der Tat die bloßen zwei Zahlen

$$\mathfrak{D}(x) = A, \quad \text{str.}\,(x) = \varepsilon_2$$

einen sehr wertvollen Anhalt zur Beschreibung eines Kollektivgegenstandes und ein Mittel zur Vergleichung von Kollektivreihen über denselben Gegenstand.

Will man einen Schritt weiter tun, so wird ε_3 in Vorzeichen und Größe Aufschluß geben über Richtung und Stärke der Asymmetrie in der Verteilung.

Auf Einzelheiten einzugehen wird sich bei der Vorführung spezieller Kollektivreihen Gelegenheit ergeben.

IX. Die zuletzt erörterten Methoden der mathematischen Beschreibung eines Kollektivgegenstandes durch Mittelwerte soll nun an einer Auswahl von Beobachtungsreihen anthropometrischer Natur zur Anwendung gebracht werden, wobei auch die Anlage der erforderlichen Rechnungen zu vollem Ausdruck kommen wird.

1. Ich beginne mit der ersten Beobachtungsreihe Quetelets, die schon in Tabelle IV mitgeteilt worden ist und sich auf die Brustumfänge schottischer Soldaten bezieht. In den ersten drei Kolonnen der Tabelle IX sind ihre Zahlen wiederholt und es schließt sich daran die Durchführung des Summenverfahrens, wobei als Ausgangswert a der letzte Argumentwert $x_{16} = 48$ angenommen ist, so daß die ganze Rechnung einheitlich geführt ist. Ihre Ergebnisse sind die folgenden:

$$S_0^- = 5\,735 + 4 = 5\,739, \quad S_0^- + z_{16} = 5\,740 = m,$$
$$S_1^- = 41\,148, \quad S_2^- = 138\,957, \quad S_3^- = 292\,698;$$

die Kolonne $s_i^{(4)}$ ist nur zur Kontrolle von S_3^- geführt; in der Tat ist $189\,154 + 103\,544 = 292\,698$.

Tabelle IX.

i	x_i engl. Zoll	z_i	$s_i^{(1)}$	$s_i^{(2)}$	$s_i^{(3)}$	$s_i^{(4)}$
1	33	3	3	3	3	3
2	34	18	21	24	27	30
3	35	81	102	126	153	183
4	36	185	287	413	566	749
5	37	420	707	1 120	1 686	2 435
6	38	749	1 456	2 576	4 262	6 697
7	39	1 075	2 531	5 107	9 369	16 066
8	40	1 079	3 610	8 717	18 086	34 152
9	41	934	4 544	13 261	31 347	65 499
10	42	658	5 202	18 463	49 810	115 309
11	43	370	5 572	24 035	73 845	189 154
12	44	92	5 664	29 699	103 544	
13	45	50	5 714	35 413	292 698	
14	46	21	5 735	138 957		
15	47	4	41 148			
16	48	1				
		5 740				

Nach den Formeln (34), worin jetzt $\delta = 1$ ist, erhält man:

$$m\eta_1 = -46\,887, \quad m\eta_2^2 = 407\,097, \quad m\eta_3^3 = -3\,117\,447,$$

daraus

$$\eta_1 = -8{,}17, \quad \eta_2^2 = 70{,}923, \quad \eta_3^3 = -617{,}639,$$

und mit Hilfe der Formeln (40)

$$\varepsilon_1 = 0, \quad \varepsilon_2^2 = 4{,}174, \quad \varepsilon_3^3 = 0{,}006.$$

Auf Grund dieser Zahlen findet man den Argumentdurchschnitt

$$A = 48 - 8{,}17 = 39{,}83 \text{ engl. Zoll}$$

und die Streuung

$$\text{str.}(x) = \varepsilon_2 = 2{,}04 \text{ engl. Zoll},$$

ins metrische Maß umgerechnet (1 engl. Zoll $= 0{,}0253998 m$):

$$A = 1{,}011 m, \quad \text{str.}(x) = 51{,}8 mm.$$

Die sechsfache Streuung beträgt etwas über 12 Zoll; innerhalb des Intervalls $34^1/_2$ bis $46^1/_2$ Zoll, das diese Ausdehnung hat, liegen 5714, also 99,55% aller Brustumfänge.

Der sehr geringe Wert von ε_3^3 weist auf einen hohen Grad der Symmetrie hin und rechtfertigt den Versuch, die Kollektivreihe durch das Gaußsche Gesetz darzustellen; das Resultat ist in der Tat ein befriedigendes. Man findet nämlich mittels ε_2 das Präzisionsmaß

$$h = \frac{1}{\varepsilon_2 \sqrt{2}} = 0{,}346;$$

die Rechnung hat sich auf die Wechselpunkte zu stützen; für die Wahrscheinlichkeit, daß ein Argumentwert zwischen die Wechselpunkte 38,5 und 41,5, die vom arithmetischen Mittel um 1,33 bezw. 1,67 abweichen, falle, ergibt sich der Ausdruck und Wert

$$\frac{1}{2}\left\{\frac{2}{\sqrt{\pi}}\int_0^{1,33\,h} e^{-t^2}\,dt + \frac{2}{\sqrt{\pi}}\int_0^{1,67\,h} e^{-t^2}\,dt\right\} = 0{,}53548,$$

welch letzterer, mit $m = 5\,740$ multipliziert, die rechnungsmäßig zu erwartende Anzahl der Argumentwerte zwischen den obigen Wechselpunkten gibt, die mit der beobachteten zu vergleichen ist. Das Ergebnis einer solchen Proberechnung ist aus dem folgenden Täfelchen zu ersehen:

Intervall	z	
	berechnet	beobachtet
38,5 — 41,5	3 083,5	3 088
36,5 — 43,5	5 235,6	5 285
34,5 — 45,5	5 697,7	5 623

2. An zweiter Stelle behandle ich ein Beobachtungsmaterial, das W. R. Macdonell im ersten Bande der monumentalen englischen Zeitschrift „Biometrika“[1]) mitgeteilt und nach anderen Methoden der Rechnung unterzogen hat. Es betrifft Messungen der Länge des linken Mittelfingers und der Körperhöhe an 3000 in englischen Strafhäusern internierten Verbrechern, die jedoch nicht zu der qualifizierten Sorte der „Gewohnheitsverbrecher“ gehören; die 3000 Fälle sind aus den umfangreichen Listen des Zentralmeßbureaus, das dem Zwecke der Identifikation von Verbrechern dient, aufs Geratewohl entnommen. Die Länge des linken Mittelfingers[2]) hat sich einerseits wegen ihrer scharfen, von der Einflußnahme des gemessenen Individuums unabhängigen Bestimmbarkeit und wegen ihrer beträchtlichen Variabilität als ein wertvolles Erkennungsmerkmal erwiesen.

Tabelle X betrifft die Mittelfingerlänge und enthält die von Millimeter zu Millimeter fortschreitende Verteilungstafel und die Summenkolonnen, die wegen der Kontrolle von S_2 bis $s^{(3)}$ geführt sind. Zum Ausgangswert ist der in der Mitte der Tafel stehende Argumentwert $x_2 = 115$ gewählt.

Die Rechnung stellt sich hier folgendermaßen:

$$
\begin{array}{lll}
S_0^+ = 1260 + 226 = 1486 & S_1^+ = 5730 & S_2^+ = 16823 \\
S_0^- = 1053 + 228 = 1281 & S_1^- = 4514 & S_2^- = 11912 \\
\Delta_0 = 205 & \Delta_1 = 1216 & \\
\Sigma_0 = 2767 & \Sigma_1 = 10244 & \Sigma_2 = 28735
\end{array}
$$

$$
\begin{array}{ll}
m = \Sigma_0 + 233 = 3000 & \\
\Delta_1 = 1216 & 2\Sigma_2 = 57470 \\
\Delta_0 = 205 & 3\Sigma_1 = 30732 \\
m\eta_1 = 1421 & \Sigma_0 = 2767 \\
\eta_1 = 0{,}474 & m\eta_2^2 = 90969 \\
 & \eta_2^2 = 30{,}323 \\
 & \eta_1^2 = 0{,}225 \\
 & \varepsilon_2^2 = 30{,}098.
\end{array}
$$

Hieraus ergeben sich die beiden Endresultate:

$$A = 115 + 0{,}474 = 115{,}474 mm$$

$$\text{str.}\,(x) = \varepsilon_2 = 5{,}48 mm.$$

---

[1]) A Journal for the Statistical Study of Biological Problems. Cambridge, 1902.

[2]) Vom Handknöchelgelenk bis zur Fingerspitze. Ihre Messung erfolgt mittels eines Schiebermaßes und ist so ausgebildet, daß die Länge innerhalb eines halben Millimeters genau festgestellt werden kann. Man vergleiche: Das anthropometrische Signalement von Alphons Bertillon. Deutsch von Dr. v. Sury. Bern, 1895.

Tabelle X.

i	x_i in mm	z_i	$s_i^{(1)}$	$s_i^{(2)}$	$s_i^{(3)}$
1	95	1	1	1	1
2	96	0	1	2	3
3	97	0	1	3	6
4	98	1	2	5	11
5	99	3	5	10	21
6	100	7	12	22	43
7	101	7	19	41	84
8	102	10	29	70	154
9	103	17	46	116	270
10	104	20	66	182	452
11	105	30	96	278	730
12	106	44	140	418	1 148
13	107	74	214	632	1 780
14	108	75	289	921	2 701
15	109	102	391	1 312	4 013
16	110	163	554	1 866	5 879
17	111	152	706	2 572	8 451
18	112	183	889	3 461	
19	113	164	1 053	11 912	
20	114	228	4 514		
21	115	233			
22	116	226	5 730		
23	117	232	1 260	16 823	
24	118	184	1 028	4 470	
25	119	162	844	3 442	12 353
26	120	163	682	2 589	8 911
27	121	126	519	1 916	6 313
28	122	91	393	1 397	4 397
29	123	89	302	1 004	3 000
30	124	44	213	702	1 996
31	125	52	169	489	1 294
32	126	35	117	320	805
33	127	31	82	203	485
34	128	25	51	121	282
35	129	7	26	70	161
36	130	8	19	44	91
37	131	2	11	25	47
38	132	6	9	14	22
39	133	2	3	5	8
40	134	0	1	2	3
41	135	1	1	1	1
		3 000			

Die sechsfache Streuung, 32,88*mm*, umfaßt tatsächlich wieder die Hauptmasse der Fälle, mit Weglassung der obersten und untersten vier Argumente 2 989 von 3 000 Exemplaren.

Tabelle XI zeigt die Verteilung der Körperlängen bei denselben 3 000 Verbrechern, in Intervallen von 1 engl. Zoll, und die Summenbildung im selben Umfang und in der gleichen Anordnung wie in Tabelle X; als Ausgangswert a ist der dichteste Argumentwert $x_{10} = 5'\ 5^1/_{16}''$ gewählt worden.

Tabelle XI.

i	x_i engl. Fuß und Zoll	z_i	$s_i^{(1)}$	$s_i^{(2)}$	$s_i^{(3)}$
1	$4'\ 8^1/_{16}''$	1	1	1	1
2	$9^1/_{16}''$	1	1	3	4
3	$10^1/_{16}''$	6	8	11	15
4	$11^1/_{16}''$	23	31	42	57
5	$5'\ {}^1/_{16}''$	48	79	121	178
6	$1^1/_{16}''$	90	169	290	468
7	$2^1/_{16}''$	175	344	634	
8	$3^1/_{16}''$	317	661	1 102	
9	$4^1/_{16}''$	393	1 295		
10	$5^1/_{16}''$	462			
11	$6^1/_{16}''$	458	2 284		
12	$7^1/_{16}''$	413	1 026	2 406	
13	$8^1/_{16}''$	264	613	1 258	
14	$9^1/_{16}''$	177	349	645	1 148
15	$10^1/_{16}''$	97	172	296	503
16	$11^1/_{16}''$	46	75	124	207
17	$6'\ {}^1/_{16}''$	17	29	49	83
18	$1^1/_{16}''$	7	12	20	34
19	$2^1/_{16}''$	4	5	8	14
20	$3^1/_{16}''$	0	1	3	6
21	$4^1/_{16}''$	0	1	2	3
22	$5^1/_{16}''$	1	1	1	1
		3 000			

Man hat jetzt in gleicher Anlage der Rechnung:

$S_0^+ = 1026 + 458 = 1484$	$S_1^+ = 2284$	$S_2^+ = 2406$
$S_0^- = 661 + 393 = 1054$	$S_1^- = 1295$	$S_2^- = 1102$
$\Delta_0 = 430$	$\Delta_1 = 989$	
$\Sigma_0 = 2538$	$\Sigma_1 = 3579$	$\Sigma_2 = 3508$

$$m = \Sigma_0 + 462 = 3000$$

$$\Delta_1 = 989$$
$$\Delta_0 = 430$$
$$m\eta_1 = 1419$$
$$\eta_1 = 0{,}473$$

$$2\Sigma_2 = 7016$$
$$3\Sigma_3 = 10737$$
$$\Sigma_0 = 2538$$
$$m\eta_2^2 = 20291$$
$$\eta_2^2 = 6{,}764$$
$$\eta_1^2 = 0{,}224$$
$$\varepsilon_2^2 = 6{,}540$$

$$A = 5'\,5{,}063'' + 0{,}473'' = 65{,}536'' = 1{,}665m$$
$$\text{str.}\,(x) = \varepsilon_2 = 2{,}55'' = 64{,}8mm.$$

Im Grunde genommen handelt es sich hier um einen Kollektivgegenstand mit zwei Argumenten: Mittelfingerlänge und Körperhöhe Die Vermessungsergebnisse sind auch in diesem Sinne tabellarisiert worden und es ergab sich eine Tafel, welche die absolute Häufigkeit jeder Verbindung aus den Mittelfingerlängen der Tabelle X und den Körperhöhen der Tabelle XI angibt und sonach das Studium des Verhältnisses dieser zwei Dimensionen gestattet. Die Tafel gibt aber auch die mittlere Länge des Mittelfingers bei jeder der in Tabelle XI angeführten Körperhöhen an und in diesem Sinne wollen wir sie benützen, um das interessante Verhalten beider Zahlen zueinander zu beleuchten. Die folgende kleine Tabelle, deren Beschreibung keine weitere Erklärung erfordert, zeigt vor allem, daß beide Dimensionen in einem nur sehr langsam veränderlichen u. zw. mit der Körperhöhe etwas steigenden Verhältnis zueinander stehen und daß die Körperhöhe $14^1/_4$- bis $14^3/_4$ mal so groß ist als die Länge des linken Mittelfingers.

3. Macdonell hat bei einer Gruppe von 1306 Verbrechern, welche der Liste der vorigen 3000 willkürlich entnommen worden waren, Kopfbreite und Kopflänge [1]) angegeben und zum Vergleiche die

[1]) Unter Kopfbreite wird die größte, zur Symmetrieebene des Körpers normale Querachse des Kopfes verstanden; ihre Endpunkte liegen in der Regel etwa zwei Finger breit ober und hinter den oberen Ohrenansätzen, fallen aber mitunter mit diesen selbst zusammen. Als Kopflänge wird die Entfernung von der Vertiefung der Nasenwurzel bis zum höchsten Punkte des Hinterkopfes bezeichnet. Beide Maße werden mit Hilfe des Kopfzirkels abgenommen und lassen sich auf einen halben Millimeter genau bestimmen. Vergl. Bertillon, l. c.

analogen Maße von 1000 Cambridger Studenten herangezogen, um zu sehen, ob sich zwischen diesen zwei gesellschaftlich und in Bezug auf Intelligenz so verschiedenen Klassen anthropometrisch ein Unterschied zu erkennen gibt.

Körperhöhe in *m*	Mittlere Länge des linken Mittelfingers in *mm*	Verhältnis $\frac{\text{Körperhöhe}}{\text{Mittelfingerlänge}}$
1,424	100,0	14,24
1,449	103,0	14,07
1,474	102,8	14,34
1,500	107,0	14,02
1,525	107,8	14,15
1,551	109,4	14,18
1,577	110,6	14,25
1,602	111,8	14,33
1,627	113,3	14,36
1,652	114,8	14,39
1,678	116,5	14,40
1,703	117,7	14,46
1,728	118,6	14,57
1,754	120,1	14,61
1,779	122,2	14,55
1,804	123,9	14,56
1,830	125,9	14,54
1 855	126,4	14,68
1,880	127,7	14,78
1,906	.	.
1,931	.	.
1,957	112,0	.

Tabelle XII enthält die primäre Verteilungstafel der Kopfbreiten, Tabelle XIII jene der Kopflängen der erstgenannten Personengruppe, beide in Intervallen von 1*mm* fortschreitend, und die zur weiteren Bearbeitung erforderliche Summenrechnung. Als Ausgangswert ist beidemale der dichteste Wert, $x_{18} = 151$ im ersten, $x_{22} = 192$ im zweiten Falle, angenommen.

Tabelle XII.

i	x_i in mm	z_i	$s_i^{(1)}$	$s_i^{(2)}$	$s_i^{(3)}$
1	134	1	1	1	1
2	135	1	2	3	4
3	136	4	6	9	13
4	137	3	9	18	31
5	138	6	15	33	64
6	139	7	22	55	119
7	140	9	31	86	205
8	141	18	49	135	340
9	142	21	70	205	545
10	143	39	109	314	859
11	144	46	155	469	1 328
12	145	65	220	689	2 017
13	146	90	310	999	3 016
14	147	92	402	1 401	4 417
15	148	86	488	1 889	
16	149	98	586	6 306	
17	150	92	2 475		
18	151	104	.		
19	152	97	1 718		
20	153	91	427	4 380	
21	154	73	336	1 291	
22	155	56	263	955	3 089
23	156	57	207	692	2 134
24	157	46	150	485	1 442
25	158	27	104	335	957
26	159	22	77	231	622
27	160	18	55	154	391
28	161	10	37	99	237
29	162	12	27	62	138
30	163	5	15	35	76
31	164	5	10	20	41
32	165	3	5	10	21
33	166	1	2	5	11
34	167	0	1	3	6
35	168	0	1	2	3
36	169	1	1	1	1
		1 306			

Tabelle XIII.

i	x_i in mm	z_i	$s_i^{(1)}$	$s_i^{(2)}$	$s_i^{(3)}$
1	171	1	1	1	1
2	172	1	2	3	4
3	173	2	4	7	11
4	174	0	4	11	22
5	175	3	7	18	40
6	176	3	10	28	68
7	177	5	15	43	111
8	178	7	22	65	176
9	179	12	34	99	275
10	180	13	47	146	421
11	181	17	64	210	631
12	182	28	92	302	933
13	183	24	116	418	1 351
14	184	43	159	577	1 928
15	185	53	212	789	2 717
16	186	57	269	1 058	3 775
17	187	55	324	1 382	5 157
18	188	68	392	1 774	6 931
19	189	83	475	2 249	.
20	190	85	560	9 180	.
21	191	96	2 809	.	.
22	192	102	.	.	.
23	193	79	2 183	.	.
24	194	83	469	6 518	.
25	195	66	386	1 714	.
26	196	66	320	1 328	4 806
27	197	56	254	1 008	3 476
28	198	43	198	754	2 468
29	199	35	155	556	1 714
30	200	30	120	401	1 158
31	201	20	90	281	757
32	202	24	70	191	476
33	203	14	46	121	285
34	204	13	32	75	164
35	205	8	19	43	89
36	206	3	11	24	46
37	207	6	8	13	22
38	208	0	2	5	9
39	209	1	2	3	4
40	210	1	1	1	1
		1 306			

Die Durchführung der Rechnungen führt zu folgenden Resultaten:

$S_0^+ = 427 + 97 = 524$	$S_1^+ = 1\,718$	$S_2^+ = 4\,380$
$S_0^- = 586 + 92 = 678$	$S_1^- = 2\,475$	$S_2^- = 6\,306$
$\Delta_0 = -154$	$\Delta_1 = -757$	$\Delta_0 = -1\,926$
$\Sigma_0 = 1\,202$	$\Sigma_1 = 4\,193$	$\Sigma_2 = 10\,686$

$$
\begin{aligned}
m = \Sigma_0 + 104 &= 1\,306 \\
\Delta_1 &= -757 \\
\Delta_0 &= -154 \\
\hline
m\,\eta_1 &= -911 \\
\eta_1 &= -0{,}7
\end{aligned}
\qquad
\begin{aligned}
2\,\Sigma_2 &= 21\,372 \\
3\,\Sigma_1 &= 12\,579 \\
\Sigma_0 &= 1\,202 \\
\hline
m\,\eta_2^2 &= 35\,153 \\
\eta_2^2 &= 26{,}91 \\
\eta_1^2 &= 0{,}49 \\
\hline
\varepsilon_2^2 &= 26{,}42
\end{aligned}
$$

$$A = 151 - 0{,}7 = 150{,}3\,mm.$$
$$\text{str.}\,(x) = \varepsilon_2 = 5{,}14\;mm.$$

Die zu Tafel XIII gehörige Rechnung lautet:

$S_0^+ = 469 + 79 = 548$	$S_1^+ = 2\,183$	$S_2^+ = 6\,518$
$S_0^- = 560 + 96 = 656$	$S_1^- = 2\,809$	$S_2^- = 9\,180$
$\Delta_0 = -118$	$\Delta_1 = -626$	
$\Sigma_0 = 1\,204$	$\Sigma_1 = 4\,992$	$\Sigma_2 = 15\,698$

$$
\begin{aligned}
m = \Sigma_0 + 102 &= 1\,306 \\
\Delta_1 &= -626 \\
\Delta_0 &= -118 \\
\hline
m\,\eta_1 &= -744 \\
\eta_1 &= -0{,}569
\end{aligned}
\qquad
\begin{aligned}
2\,\Sigma_2 &= 31\,396 \\
3\,\Sigma_1 &= 14\,976 \\
\Sigma_0 &= 1\,204 \\
\hline
m\,\eta_2^2 &= 47\,576 \\
\eta_2^2 &= 36{,}44 \\
\eta_1^2 &= 0{,}32 \\
\hline
\varepsilon_2^2 &= 36{,}12
\end{aligned}
$$

$$A = 192 - 0{,}569 = 191{,}431\,mm$$
$$\text{str.}\,(x) = \varepsilon_2 = 6{,}02\,mm.$$

Die beiden vorstehenden Reihen lassen eine befriedigende Darstellung durch das **Gaußsche** Gesetz zu, wie die nachfolgende Tabelle erkennen läßt; die Präzisionsmaße, aus den mittleren Fehlern ε_2 berechnet, sind dabei

$$\frac{1}{5{,}14\sqrt{2}} = 0{,}14 \text{ bezw. } \frac{1}{6{,}02\sqrt{2}} = 0{,}12.$$

Kopfbreite			Kopflänge		
Intervall *mm*	z berechnet	z beobachtet	Intervall *mm*	z berechnet	z beobachtet
145,5 — 155,5	884,6	879	186,5 — 196,5	788,5	783
140,5 — 160,5	1 243,4	1 238	181,5 — 201,5	1 188,9	1 172
135,5 — 165,5	1 302,0	1 302	176,5 — 206,5	1 291,7	1 288

Die Tabellen XIV und XV betreffen die Kopfmessungen an 1000 Studenten der Universität Cambridge; die Verteilungstafeln schreiten hier nach einem größeren Intervall, von $^1/_{10}$ zu $^1/_{10}$ engl. Zoll, d. i. etwa $2^1/_2$*mm*, fort und zeigen demgemäß im einzelnen geringere Unregelmäßigkeiten als die früheren. Die Rechnung ist hier wieder einheitlich geführt, wobei der größte Argumentwert zum Ausgangswert a genommen ist. Nicht übersehen darf man den Umstand, daß nun δ nicht 1, wie in allen bisherigen Fällen, sondern 0,1 ist. Tabelle XIV bezieht sich auf die Kopfbreiten, Tabelle XV auf die Kopflängen.

Tabelle XIV.

i	x_i engl. Zoll	z_i	$s_i^{(1)}$	$s_i^{(2)}$	$s_i^{(3)}$
1	5,5	3	3	3	3
2	5,6	12	15	18	21
3	5,7	43	58	76	97
4	5,8	80	138	214	311
5	5,9	131	269	483	794
6	6,0	236	505	988	1 782
7	6,1	185	690	1 678	3 460
8	6,2	142	832	2 510	5 970
9	6,3	99	931	3 441	9 411
10	6,4	37	968	4 409	13 820
11	6,5	15	983	5 392	
12	6,6	12	995	19 212	
13	6,7	3	6 387		
14	6,8	2			
		1 000			

Die Rechnung gestaltet sich wie folgt:

$S_0^- = 995 + 3 = 998$ $\qquad S_1^- = 6\,387$ $\qquad S_2^- = 19\,212$

$m = S_0^- + 2 = 1000$

$$\begin{aligned} S_1^- &= 6\,387 \\ S_0^- &= 998 \\ \hline m\,\eta_1 &= -7\,385\,.\,0{,}1 \\ \eta_1 &= -0{,}7385 \end{aligned} \qquad \begin{aligned} 2\,S_2^- &= 38\,424 \\ 3\,S_1^- &= 19\,161 \\ S_0^- &= 998 \\ \hline m\,\eta_2^2 &= 58\,583\,.\,0{,}01 \\ \eta_2^2 &= 0{,}5\,858 \\ \eta_1^2 &= 0{,}5\,454 \\ \hline \varepsilon_2^2 &= 0{,}0\,404 \end{aligned}$$

$A = 6{,}8 - 0{,}7\,385 = 6{,}0\,615$ engl. Zoll $= 153{,}95\,mm$,
str. $(x) = \varepsilon_2 = 0{,}201$ engl. Zoll $= 5{,}11\,mm$.

Tabelle XV.

i	x_i engl. Zoll	z_i	$s_i^{(1)}$	$s_i^{(2)}$	$s_i^{(3)}$
1	6,8	1	1	1	1
2	6,9	2	3	4	5
3	7,0	6	9	13	18
4	7,1	21	30	43	61
5	7,2	43	73	116	177
6	7,3	58	131	247	424
7	7,4	100	231	478	902
8	7,5	145	376	854	1 756
9	7,6	175	551	1 405	3 161
10	7,7	180	731	2 136	5 297
11	7,8	98	829	2 965	8 262
12	7,9	75	904	3 869	12 131
13	8,0	61	965	4 834	16 965
14	8,1	22	987	5 821	22 786
15	8,2	9	996	6 817	29 603
16	8,3	3	999	7 816	
17	8,4	—	999	37 419	
18	8,5	—	8 815		
19	8,6	1			
		1 000			

Die zugehörige Rechnung lautet:

$S_0^- = 999 + 0 = 999 \qquad S_1^- = 8\,815 \qquad S_2^- = 37\,419$

$m = S_0^- + 1 = 1\,000$

$S_1^- = 8\,815$
$S_0^- = 999$
$m\eta_1 = -9\,814 . 0{,}1$
$\eta_1 = -0{,}9814$

$2S_2^- = 74\,838$
$3S_1^- = 26\,445$
$S_0^- = 999$
$m\eta_2^2 = 102\,282 . 0{,}01$
$\eta_2^2 = 1{,}0228$
$\eta_1^2 = 0{,}9631$
$\varepsilon_2^2 = 0{,}0597$

$$A = 8{,}6 - 0{,}9814 = 7{,}6186 \text{ engl. Zoll} = 193{,}48\,mm$$
$$\text{str.}\,(x) = \varepsilon_2 = 0{,}244 \text{ engl. Zoll} = 6{,}19\,mm.$$

Die Zusammenstellung der auf die beiden Personenkategorien bezüglichen Resultate gibt folgendes Bild:

Kategorie	Kopfbreite		Kopflänge		Verhältnis Länge/Breite
	A	ε	A	ε	
Verbrecher	150,3	5,14	191,4	6,02	1,237
Studenten	153,9	5,11	193,5	6,19	1,257

Es geht daraus ein unverkennbarer Unterschied in den Kopfdimensionen hervor; beide sind im Durchschnitt bei den Studenten größer, die Kopfbreite um 3,6, die Länge um 2,1 *mm*. Ihr Verhältnis variiert von der einen Kategorie zur anderen nur unerheblich und weist bei den Studenten auf eine relativ größere Kopfbreite hin. Die Streuung ist bei den Kopflängen beträchtlich größer als bei den Breiten, jene erstrecken sich über ein ausgedehnteres Intervall als diese. Zum Vergleiche seien auch die bezüglichen Angaben Bertillons (l. c.) angeführt: er bezeichnet 156 *mm* als mittlere Kopfbreite und bemerkt, daß sie selten unter 140 und selten über 169 *mm* beträgt; die mittlere Kopflänge beziffert er mit 187 *mm* (hienach wäre Kopflänge : Kopfbreite = 1,2) und sagt, daß sie selten unter 170 *mm* und selten über 200 *mm* falle. Es darf nicht übersehen werden, daß sich diese Angaben auf ein anderes Material beziehen und kaum das Resultat einer sorgfältigen Rechnung sein dürften.

4. Zum Schlusse führe ich ein derselben englischen Zeitschrift entlehntes Beobachtungsmaterial vor und teile die Resultate meiner Bearbeitung desselben nach der nämlichen Methode mit; es handelt sich um Körperhöhemessungen auf verschiedenen Altersstufen, die A. O. Powys

an Verbrechern in Neusüdwales (Australien) ausgeführt hat. Die Aufmerksamkeit richtet sich dabei hauptsächlich auf die Änderungen, welche die Verteilung der Körperhöhen mit fortschreitendem Alter erfährt. Tabelle XVI bringt lediglich die Verteilungstafeln für die Altersquinquennien 20—25, 25—30, dann die Altersdezennien 30—40, 40—50, 50—60 und schließlich für die restlichen Alter, nach Intervallen von 1 engl. Zoll fortschreitend; der jeweilige Umfang ist am Fuße der Tafel ersichtlich gemacht. Die Summenrechnung ist unterdrückt, um nicht ein allzugroßes Ziffernmaterial bringen zu müssen; die in Tabelle XVII mitgeteilten Resultate ermöglichen es, ihre etwaige selbständige Durchführung zu kontrollieren.

Tabelle XVI.

x_i engl. Maß	20 — 25	25 — 30	30 — 40	40 — 50	50 — 60	über 60
	z_i	z_i	z_i	z_i	z_i	z_i
unter 4′ 7″	1	3	2	7	3	
4′ 7½″	—	3	1	—	—	2
8½	2	1	2	2	—	2
9½″	3	2	3	1	2	1
10½″	—	9	6	12	7	3
11½″	12	25	27	16	8	25
5′ ½″	42	70	99	52	35	44
1½″	75	164	209	156	117	80
2½″	158	323	462	363	157	164
3½″	275	633	852	567	301	230
4½″	551	1 056	1 335	867	481	338
5½″	725	1 427	1 793	1 154	569	411
6½″	**939**	2 001	2 398	1 391	**682**	**441**
7½″	882	**2 019**	**2 406**	**1 431**	655	384
8½″	734	1 826	2 103	1 153	568	297
9½″	482	1 410	1 351	807	363	228
10½″	281	876	1 027	520	232	123
11½″	151	455	568	263	121	49
6′ ½″	82	287	300	142	65	25
1½″	35	105	122	61	30	11
2½″	8	37	33	27	10	2
3½″	3	11	11	8	1	2
4½″	1	—	5	1		
5½″		2	1	2		
6½″		1	1			
7½″		—				
8½″		1				
$m =$	5 442	12 747	15 117	9 003	4 407	2 862

Tabelle XVII.

Altersintervall	Englisches Maß		Metermaß	
	A Zoll	ε_2 Zoll	A m	ε_2 mm
20 — 25	66,952	2,437	1,699	61,8
25 — 30	67,305	2 547	1,708	65,4
30 — 40	67,153	2,588	1,704	65,7
40 — 50	66,911	2,917	1,698	66,4
50 — 60	66,745	2,631	1,694	66,8
über 60	66,256	2,667	1,681	67,6

Bei der Betrachtung der Tabelle XVI ist Vorsicht geboten, weil die Umfänge der einzelnen Verteilungsreihen sehr verschieden sind — der kleinste Umfang beträgt etwa $^1/_5$ des größten —; der bloße Anblick könnte zu dem Schlusse verleiten, daß die Ausbreitung der Verteilungsreihe vom Alter 30 aufwärts abnehme. Hingegen sind die Zahlen der Tabelle XVII von den Umfängen unabhängig und daher unmittelbar vergleichbar.

Die dichtesten Werte, die in Tabelle XVI durch stärkeren Druck hervorgehoben sind, zeigen einen analogen Gang wie die arithmetischen Mittel; bis auf eine Ausnahme — die Altersstufe 40—50 — fällt dieses in das Intervall des dichtesten Argumentwertes, in dem Ausnahmefalle in das darüber liegende. Die mittlere Körperhöhe erreicht zwischen 25 und 30 Jahren ihren größten Wert und nimmt von da an allmählich ab. Bemerkenswert ist die ständige Zunahme der Streuung, also die abnehmende Konzentration der Verteilung um das arithmetische Mittel.

X. Die weitestgehende Forderung, die man an die mathematische Beschreibung eines Kollektivgegenstandes stellen kann, geht dahin, seine Verteilung in ihrem ganzen Zuge durch eine analytische Formel darzustellen, ohne über die Gestalt dieser Formel eine Voraussetzung zu treffen Die Darstellung kann sich auf die Verteilungsfunktion selbst richten; aus früher angegebenen Gründen wird es jedoch vorteilhafter sein, die Summenfunktion zum Ausgangspunkt zu nehmen.

Dieser Forderung kann nur durch eine willkürliche Funktion genügt werden, weil nur eine solche fähig ist, allen Modalitäten der Verteilung, die bei Kollektivgegenständen jeder denkbaren Provenienz auftreten können, angepaßt zu werden. Eine derartige Funktion ist nur durch einen unendlichen Prozeß, also insbesondere durch eine unendliche Reihe,

darstellbar; ist aber die Form dieser Reihe zweckmäßig gewählt, in dem Sinne, daß bei ihrer numerischen Auswertung auf der Basis einer gegebenen Verteilungstafel die Glieder rasch abnehmen, so wird schon ein kurzes Bruchstück der Reihe ausreichen, um die Verteilung innerhalb solcher Grenzen formelmäßig wiederzugeben, die der Natur der Sache entsprechen.

Eine solche analytische Darstellung einer willkürlichen Funktion bildet den Kernpunkt einer Schrift[1]), welche H. Bruns im Vorjahre veröffentlicht hat. Die Glieder der Reihe, die er für die Summenfunktion $\mathfrak{S}(X)$, unmittelbar für die Funktion $2\,\mathfrak{S}(X) - 1$, gegeben, schreiten nach der Funktion

$$\Phi(X) = \frac{2}{\sqrt{\pi}} \int_0^X e^{-t^2} dt$$

und ihren Ableitungen $\Phi'(X)$, $\Phi''(X)$, ... fort, und ihre Koeffizienten setzen sich linear aus den Potenzmitteln aller Grade zusammen.

Der Erfolg dieser Formel bei ihren bisherigen Anwendungen auf verschiedenartige Kollektivreihen ist in ihrer Anlehnung an das Gaußsche Exponentialgesetz zu erblicken, das doch gewissermaßen den Grundzug fast aller bislang untersuchten Kollektivgegenstände bildet; dadurch war eine rasche Konvergenz der Reihe von vornherein gesichert. Die große Tragweite der Formel besteht aber darin, daß sie die Abweichungen von jenem Gesetze, das vordem als die durchgehende Norm angesehen wurde, in systematischer Weise wiederzugeben gestattet.

Für die Funktion $\Phi(X)$ existieren Tabellen schon seit langem; für ihre Ableitungen hat Bruns solche berechnet und in dem zitierten Buche veröffentlicht. Das Schwergewicht der Arbeit fällt in die Berechnung der Koeffizienten; da aber diese sich aus den Potenzmitteln der Abweichungen vom Argumentdurchschnitt zusammensetzen, so ist im vorigen der ganze Rechenapparat auch für die Aufstellung der Brunsschen Φ-Reihe enthalten.

Bezüglich der Aufstellung dieser Reihe selbst muß ich jedoch auf das genannte Werk verweisen oder auf die Darstellung, die ich dafür in einer Abhandlung[2]) gegeben, die zugleich eine umfangreiche Anwendung der Reihe auf ein wichtiges biometrisches Problem, auf die Verteilung der Gestorbenen nach dem Alter, enthält.

---

[1]) Bereits unter III. zitiert.

[2]) Mitteilungen des „Österreichisch-ungarischen Verbandes der Privat-Versicherungs-Anstalten", Neue Folge, 3. Band, 1. Heft (1907).